

```
*      124231 124330: gap of      100 bp
*      124331 159550: contig of 35220 bp in length.
FEATURES
    source
        1..159550
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11"
            /clone="RP11-278E23"
            /clone_1fb="RPC1-11 Human Male BAC"
            1..1729
                misc_feature
                    /note="assembly_fragment"
                    1830..3716
                        misc_feature
                            /note="assembly_fragment"
                            3817..6050
                                misc_feature
                                    /note="assembly_fragment"
                                    7051..9136
                                        misc_feature
                                            /note="assembly_fragment"
                                            9237..17859
                                                misc_feature
                                                    /note="assembly_fragment"
                                                    17960..26939
                                                        misc_feature
                                                            /note="assembly_fragment"
                                                            27040..36269
                                                                misc_feature
                                                                    /note="assembly_fragment"
                                                                    36370..48574
                                                                        misc_feature
                                                                            /note="assembly_fragment"
                                                                            48675..62315
                                                                                misc_feature
                                                                                    /note="assembly_fragment"
                                                                                    62416..77457
                                                                                        misc_feature
                                                                                            /note="assembly_fragment
                                                                                            clone_end:SP6
                                                                                            vector_side:right"
                                                                                            77558..96884
                                                                                                misc_feature
                                                                                                    /note="assembly_fragment"
                                                                                                    96985..124230
                                                                                                        misc_feature
                                                                                                            /note="assembly_fragment"
                                                                                                            124331..159550
                                                                                                                misc_feature
                                                                                                                    /note="assembly_fragment
                                                                                                                    clone_end:T7
                                                                                                                    vector_side:right"
BASE COUNT      42186 a 37364 c 37475 g 41307 t 1218 others
ORIGIN
Query Match      70.0%: Score 349.8; DB 2; Length 159550;
Best Local Similarity 85.1%: Pred. No. 7.6e-24;
Matches 416; Conservative 0; Mismatches 67; Indels 6; Gaps 2;
QY      18 GAGGACGCTGCCCCAGGGCATCGCTAATGAGGAGCGCCGACGAGGGCATCGCTAATGAGGAC 77
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124879 GAGGACGCGCCGCCACGCGCATCGCCAGAGAGAGCGCCGCCACGCGCATCGCCAGCGAGGAC 124820
QY      78 ACCACCCAGTCGATCGCCAGCAGGAAAGCCGCCACGAGGCATCGC--CGAGAGCCGCAATC 134
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124819 GCGCGCCAGGGGATCGGCACGAGAGAGCGCCGCCACGAGGCATCGCCAGCAGGACGCGCC 124760
QY      135 CAGGGCATCGCCACAGGAGGAGGTGCCCCAGGGCATCGCCCAATGGGGTGGCGGACGAGGCG 194
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124759 CAGGGCATCGCCACAGAGAGAGCGCCGCCACGAGGCATCGCCCAACGAGAGAGCGCCGCCAGG 124700
QY      195 ATTCGCAATGAGAGCGCCACGAGGAGTCGCAACTGGAGCGCCGTCACGAGCTTCGCGC 254
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124699 ATCGCCACAGAGGGCGCGCCGAGGAGCATCGCCCAAGAGGAGCGCCGCCAGGGCATCGCC 124640
QY      255 AACGGGAGCGCCCTCTCACTTCGCGCAAGGGAGCGCCGCCAGGAGCATCGCCCAACGGG 314
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124639 AAGAGAGGGCGCCGCCACGCGCATCGCCAAACAGAGAGCGCCGCCAGGGCATCGCCCAACGAG 124580
QY      315 GAGCGCACCAAGGGCATGGGCAACGAGGTGTCACATCCAGGGGATGGCTAAGAGAGCGGC 374
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124579 GACTCCGCCACGGAATCGCCAGGAGAGCGCCGCCACGAGCATCGCCAGCGAGGAGCGCC 124520
QY      375 GTCCAGGGGATCGCTAAGAGAGGTGGCGCCGAGGAGCATCGCCCAACGAGAGCGCCGCCAG 434
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 124519 GTCCAGGGGATCGCCACAGGAGAGCGCCGCCAGGAGCATCGCCCAACGAGAGCGCCGCCAG 124460
QY      435 GGAATCGCC--GAGGATTCGCACAGGGGATCGCCCAACGAGAGCGCCGCCAGGAGCATC 491
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124459 GGCATCGCCCAAGGAGAGCGCCGCCAGGAGCATCGCCCAAGAGAGCGCCGCCAGGAGCATC 124400
QY      492 GCCAACAG 500
      ||||||| ||
Db 124399 GCCAACGAG 124391

RESULT 11
AC008103/c      98557 bp      DNA      linear      PRI 27-MAY-2000
LOCUS      Homo sapiens chromosome 22q12 clone pac699j1, complete sequence.
DEFINITION
AC008103
VERSION
AC008103.27      GI:7229751
KEYWORDS
SOURCE
ORGANISM
    human.
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
    1 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Homo sapiens Chromosome 22q12 PAC Clone pac699j1 In CES-DGCR Region
    unpublished
    2 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (22-JUL-1999) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    3 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (08-DEC-1999) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    4 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (07-DEC-1999) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    5 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (08-DEC-1999) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    6 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (17-DEC-1999) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    7 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (26-JAN-2000) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    8 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (01-FEB-2000) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
    OK 73019, USA
    9 (bases 1 to 98557)
    Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
    Direct Submission
    Submitted (03-FEB-2000) Department Of Chemistry And Biochemistry,
    The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
```

```
REFERENCE
AUTHORS      OK 73019, USA
TITLE         10 (bases 1 to 98557)
JOURNAL       Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
              Direct Submission
              Submitted (05-FEB-2000) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              11 (bases 1 to 98557)
REFERENCE     Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
AUTHORS      Direct Submission
TITLE        Submitted (08-FEB-2000) Department Of Chemistry And Biochemistry,
JOURNAL      The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              12 (bases 1 to 98557)
REFERENCE     Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
AUTHORS      Direct Submission
TITLE        Submitted (10-FEB-2000) Department Of Chemistry And Biochemistry,
JOURNAL      The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              13 (bases 1 to 98557)
REFERENCE     Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
AUTHORS      Direct Submission
TITLE        Submitted (14-FEB-2000) Department Of Chemistry And Biochemistry,
JOURNAL      The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              14 (bases 1 to 98557)
REFERENCE     Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
AUTHORS      Direct Submission
TITLE        Submitted (12-MAR-2000) Department Of Chemistry And Biochemistry,
JOURNAL      The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              15 (bases 1 to 98557)
REFERENCE     Wang,Q., Fang,F., Fu,Y., Pan,H., Edelman,L. and Roe,B.A.
AUTHORS      Direct Submission
TITLE        Submitted (27-MAY-2000) Department Of Chemistry And Biochemistry,
JOURNAL      The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
              On Mar 12, 2000 this sequence version replaced gi:6958024.
              Because these overlapping clones came from different libraries
              there are numerous instances of insertions, deletions, and single
              nucleotide polymorphisms in the overlapping regions below.
              AC008132(pac95506)136107 145670 (0) overlaps AC008103(pac699j1) 1
              9536 (89001) AC007981(b291) 132871 175358 (0) overlaps
              AC008103(pac699j1) 1 44355 (54202) AC007325(b48) 71505 165050 (0)
              overlaps AC008103(pac699j1) 1 94144 (4413) AC008103(pac699j1) 56194
              98557 (0) overlaps AC007326(p423) 1 42365 (60533).
FEATURES
source       1..98557
              location/Qualifiers
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="22q12"
              /clone="pac699j1"
BASE COUNT   25600 a 24617 c 24180 g 24160 t
ORIGIN
Query Match      69.6%; Score 348.2; DB 9; Length 98557;
Best Local Similarity 84.9%; Pred. No. 1.2e-23;
Matches 415; Conservative 0; Mismatches 68; Indels 6; Gaps 2;

QY 18 GAGAGAGCGCCAGGCGATCGCTAATGAGGACGCCGACGAGCGATCGCTAATGAGGAC 77
      ||||||| ||||| || ||||| || ||||||| || ||||||| || |||||||
Db 27916 GAGGACGCGCCGACGGAATCGGCAGGAGAGCGCCGCCGACGCGATCGGCACGACG 27857
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 78 ACCACGACATGATGCCCAAGAGGAGGAGGAGGCGCCAGGCGATCGGC---GAGGACGCAATC 134
      || ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 27856 GCCGCCAGGCGATCGCCACGCGAGAGCGCGCCCAAGGCGATCGCCAGGAGAGAGCGCCGCC 27797
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 135 CAGGCGATCGCCACGAGAGAGGTTGCCACGAGGATCGCCCAATGGGGTGCAGCCACAGGCG 194
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 27796 CAGGCGATCGCCACGAGAGAGCGCGCCAGGCGATCGCCCAAGAGAGAGCGCGCCAGGCG 27737
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 195 ATCGGCATGAGGAGCGCCACGAGGCGATCGCCACTGGGAGCGCGCTTCGACGAGGCTTGGC 254
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```
Db 27736 ATCGCCAAAGAGAGACGCCGCCACAGGCGATCGCCAAAGAGAGACTCCGCCACGGAATCGCC 27677
QY 255 AAGGGGAGACCCCGTCCTCAGCTTGGCCCAACGGGAGAGCGCCGCCAGGCGATCGCCCAAGGG 314
      || ||||||| ||||| || ||||||| || ||||||| || ||||||| || ||||||| || |||||||
Db 27676 AGCGAGACGCCCGCCACGCGATCGCCACGAGAGCGCGTCCTCAGGCGATCGCCCAAGGAG 27617
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 315 GAGCCACCAAGGGCGATGGGCGCAAGAGAGTCCACATCCACGCGATCGCTAACGAGAGCGCC 374
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 27616 GAGCGCCGCCAGGCGATCGCCACAGAGAGAGCGCCGCCAGGCGATCGCCCAAGAGAGAGCGCC 27557
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 375 GTCCAGGCGATCGCTTAACGAGGTGGCGCCCGAGGCGATCGCCACAGAGAGCGCCCGCCAG 434
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 27556 GCCAGGCGATCGCCCAAGAGAGAGGAGCGCGCCCGAGGCGATCGCCCAAGAGAGCGCCCGCCAG 27497
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 435 GGAATGGCC---GAGATGTGCGACAGGGCGATCGCCCAACGAGAGAGCGCGCCAGGCGATC 491
      || ||||||| ||||| || ||||||| || ||||||| || ||||||| || |||||||
Db 27496 GGCATCGCCCAAGAGAGAGCGCGCCCGAGGCGATCGCCCAACGAGAGAGCGCCCGCCAGGCGATC 27437
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 492 GCCACACAG 500
      ||||| ||
Db 27436 GCCACGAG 27428

RESULT 12
AC007708/c 106650 bp DNA linear PRI 31-MAY-2000
LOCUS      Homo sapiens chromosome 22q11 clone b563b9, complete sequence.
DEFINITION AC007708
ACCESSION  AC007708
VERSION     AC007708.13 GI:5923689
KEYWORDS   HTG.
SOURCE      human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 106650)
AUTHORS     Zhan,M. and Roe,B.A.
TITLE       Homo sapiens Chromosome 22q11 BAC Clone b563b9 In BCRU2-GGT Region
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 106650)
AUTHORS     Zhan,M. and Roe,B.A.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
            3 (bases 1 to 106650)
REFERENCE   Zhan,M. and Roe,B.A.
AUTHORS     Direct Submission
TITLE       Submitted (25-SEP-1999) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
            4 (bases 1 to 106650)
REFERENCE   Zhan,M. and Roe,B.A.
AUTHORS     Direct Submission
TITLE       Submitted (13-MAY-2000) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
            5 (bases 1 to 106650)
REFERENCE   Zhan,M. and Roe,B.A.
AUTHORS     Direct Submission
TITLE       Submitted (31-MAY-2000) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
            On Sep 25, 1999 this sequence version replaced gi:5919292.
            Because these overlapping clones came from different libraries
            there are numerous instances of insertions, deletions, and single
            nucleotide polymorphisms in the overlapping regions below.
            AC008018(b379n11) 146090 180888 (0) overlaps AC007708(b563b9) 1
            34927 (71723) AC007708(b563b9) 1843 106650 (0) overlaps
            AC009288(p413m7) 1 101851 (39025).
FEATURES
source       1..106650
              location/Qualifiers
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
```


Query Match	Best Local Similarity	Matches	414; Conservative	0; Mismatches	69; Indels	6; Gaps	2;
18	GAGAGCGTCCAGGCGATCGCTAATGAGAGCCGACAGGAGCATGCTAATGAGAC	77					
80846	GAGAGCGCGGCCAGGCGATCGCCAGTAGAGAGCGCCGACAGGCGATCGCCAGAGAGAC	80787					
QY	78	ACACCCAGTGCATGCGCCACAGAGAGCGCCGACAGGCGATCGC--CGAGAGCGGCATC	134				
Db	80786	GCGCGCCAGCATGCGATCGCCAGAGAGAGCGCCGACAGGCGATCGCAGAGAGCGGCATC	80727				
QY	135	CAGGCGATCGCCACAGAGAGAGTGTCCAGAGGCGATCGCCAGTGGGTGCGCCGACAGGCG	194				
Db	80726	CAGGCGATCGCCACAGAGAGAGTGTCCAGAGGCGATCGCCAGTGGGTGCGCCGACAGGCG	80667				
QY	195	ATCGCCAAATGAGAGAGCGCACCCAGGCGATCGCCAGTGGGTGCGCCGACAGGCGATCGCC	254				
Db	80666	ATCGCCAAATGAGAGAGCGCACCCAGGCGATCGCCAGTGGGTGCGCCGACAGGCGATCGCC	80607				
QY	255	AACGGGAGCGCGTCTCAGCTTCGCCAACGGGAGCGCGCCAGGCGATCGCCAGAGCGG	314				
Db	80606	AACGGGAGCGCGCGCCAGGCGATCGCCAGAGAGCGCGCGCCAGGCGATCGCCAGAGAGG	80547				
QY	315	GAGCGCCACCAAGGGATGGGCGAACGAGGTGACATCCAGCGGCGATCGCTACAGAGAGCGCG	374				
Db	80546	GAGCGCGCCAGGCGATCGCTACAGAGAGCGCGCGCCAGAGAGGCGATCGCCAGAGAGTGGC	80487				
QY	375	GTCACAGGCGATCGCTACAGAGAGTGGCGCGCCAGGCGATCGCCAGAGAGAGCGCGCGCAG	434				
Db	80486	GCCACAGGCGATCGCCAGAGAGAGCGCGCGCCAGGCGATCGCCAGAGAGAGCGCGCGCAG	80427				
QY	435	GGAATGCGC---GAGGATGTGCGACAGAGGCGATCGCCAGAGAGAGCGCGCGCGCAGGCGATC	491				
Db	80426	GCGATGCGCCAAAGAGAGAGC	80367				
QY	492	GCCACAGAG	500				
Db	80366	GCCACAGAG	80358				
RESULT 13	AC009288/c	140876 bp	DNA	linear	PRI 31-MAY-2000		
LOCUS	AC009288	140876 bp	DNA	linear	PRI 31-MAY-2000		
DEFINITION	Human sapiens chromosome 22q11 clone p413m7, complete sequence.						
ACCESSION	AC009288						
VERSION	AC009288.13						
KEYWORDS	HTG.						
SOURCE	human.						
ORGANISM	Human sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
AUTHORS	1 (bases 1 to 140876)						
TITLE	Hua, A., Emanuel, B. and Roe, B. A.						
JOURNAL	Human sapiens Chromosome 22q11 PMC Clone p413m7 In BCR12-GST Region						
REFERENCE	unpublished						
AUTHORS	2 (bases 1 to 140876)						
TITLE	Hua, A., Emanuel, B. and Roe, B. A.						
JOURNAL	Direct Submission						
REFERENCE	Submitted (13-AUG-1999) Department Of Chemistry And Biochemistry,						
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,						
TITLE	OK 73019, USA						
JOURNAL	3 (bases 1 to 140876)						
REFERENCE	Hua, A., Emanuel, B. and Roe, B. A.						
AUTHORS	Direct Submission						
TITLE	Submitted (13-NOV-1999) Department Of Chemistry And Biochemistry,						
JOURNAL	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,						
REFERENCE	OK 73019, USA						

REFERENCE	AUTHORS	TITLE	JOURNAL
4 (bases 1 to 140876)	Hua, A., Emanuel, B. and Roe, B. A.	Direct Submission	Submitted (13-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
5 (bases 1 to 140876)	Hua, A., Emanuel, B. and Roe, B. A.	Direct Submission	Submitted (31-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT	On May 13, 2000 this sequence version replaced gi:6056212.		
FEATURES	Location/Qualifiers		
SOURCE	1..140876		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="22q11"		
	/clone="P413m7"		
BASE COUNT	38395 a	33854 c	32663 g 35964 t
ORIGIN			
Query Match	69.0%	Score 345;	DB 9; Length 140876;
Best Local Similarity	84.5%	Pred. No. 2,je-23;	
Matches 413;	Conservative 0;	Mismatches 70;	Indels 6; Gaps 2;
OY 18	GAGGACGCTGCCCGAGGGCATCGCTAATAGAGACCGCCAGCGGCATCGCTAATAGAGAC	77	
Db 78832	GAGATGATCCCTCCCGAGGGCATCGCCAAAGAGAGAGCGCCCGCCGATCGCCAAAGAGAGAC	78773	
OY 78	ACGACCCGATGCTGCTCCCAAGAGAGAGCGCCCGCCAGGGCATCGC--CGAGACGCGCATTC	134	
Db 78772	GCGCGCCAGGGGCGATCGCCAAAGAGAGAGCGCCCGCCAGGGCATCGCTAAGAGAGCGCCGCGAC	78713	
OY 135	CAGGGCATCGCCACAGAGAGGTGCCCGAGGGCATCGCCATGGGGTTCGCCGACAGGGC	194	
Db 78712	AAGGGACCGCCACAGAGAGATCGCCCGCCAGCGCATCGCCAAAGAGAGCGCCCGCCAGCGC	78653	
OY 195	ATGCCCATGAGAGAGCGCACCGAGGGCATCGGCAATGGGAGCGCGTTCACAGGGCTTCGCC	254	
Db 78652	ATGCCCGAGGAGACCGCCCGCCACGGAATCGCCAGCGAGAGCGCGCCCGCCAGGGCATCGCC	78593	
OY 255	AACGGGAGACCGCTCTCTAGCTTCGCTCAAGGGGAGCGCGCCCGAGGGCATCGCCCAACGGG	314	
Db 78592	AAGGAGAGCGCCCGCCAGGGGACCGCGCAACGAGAGAGCGCGCCCGCCAGGGCATCGCCCAACGGG	78533	
OY 315	GAGCGCCACCAAGGGGATGGGCAACAGGTCACTCCACGGGCATCGCTAAGAGAGAGCGCC	374	
Db 78532	GAGCGCGCCCGACGGCATTCGCGACGAGAGCGCGCCCGCGGCATCGCCACGAGAGAGCGCC	78473	
OY 375	GTCGAGGGCATCGCTACGAGGTGGCGCCCGCCCAAGGGCATCGCCCAACGAAGAGCGCGCCAG	434	
Db 78472	GCCGAGGGCATCGCCCAACGAGAGACCGCCCGCCAGGGCATCGCCCAACGAGAGAGCGCGCTTCG	78413	
OY 435	GGATATGCC--CGAGGATGTCGACAGGGCATCGCCCAACGAGAGCGCGCCCGAGGGCATTC	491	
Db 78412	GCGATCGCCCAACAGAGAGAGCGCCCGCCAGGGCATCGCCCAACGAGAGAGCGCGCCCGAGGGCATTC	78353	
OY 492	GCCCAACAG 500		
Db 78352	GCCCAACGAG 78344		
RESULT 14	AC007981	17538 bp	DNA linear PRI 27-MAY-2000
LOCUS	AC007981/c		
DEFINITION	Homo sapiens chromosome 22q11 clone b291, complete sequence.		
ACCESSION	AC007981		
VERSION	AC007981.46	GI:7958972	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 175358)
TITLE	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
JOURNAL	Homo sapiens Chromosome 22q11 BAC Clone b291 In CES-DGCR Region Unpublished
REFERENCE	2 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (07-JUL-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	3 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (08-JAN-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	4 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (12-JAN-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	5 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (15-JAN-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	6 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	7 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	8 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	9 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (20-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	10 (bases 1 to 175358)
AUTHORS	Fu,Y., Pan,H., Shalkh,T., Kurahashi,H., Emanuel,B. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (27-MAY-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT	On May 20, 2000 this sequence version replaced g1:7940350. Because these overlapping clones came from different libraries there are numerous instances of insertions, deletions, and single nucleotide polymorphisms in the overlapping regions below. AC008132(pacg9506) 4747 145670 (0) overlaps AC007981(b291) 1 1424255 (33936) AC007981(b291) 59901 175358 (0) overlaps AC007325(b48) 1 115948 (49102) AC007981(b291) 138871 175358 (0) overlaps AC008103(pacg99j1) 1 44355 (54202).
FEATURES	Location/Qualifiers 1..175358 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="22q11" /clone="b291"
SOURCE	

BASE COUNT	47856 a	42018 c	40291 g	45193 t
Query Match	68.8%; Score 344;	DB 9;	Length 175358;	
Best Local Similarity	84.3%; Pred. No. 2,4e-23;			
Matches	413; Conservative	0; Mismatches	70; Indels	7; Gaps
				2;
OY	18	GAGGACCGTGGCCAGGCGATCGCTAATATAGAACGCCGCACAGGGCATGTGTAATAGAGAC	77	
Db	160734	GAGGACGCGCCCAAGGCGATGCCACGACGAGACGCGCCACGGATGCGCAGCGAGGAC	160695	
OY	78	AACCACCGATCGATCGCCCAAGAGGAAGCCGCCAGGGCATTCG--CGAGAGCGCCATC	134	
Db	160694	GCCCCCGAGGCGATGCCACGAGAGGACGCCGCCAGGGCATTCGCCAGGAGAGCGCCC	160635	
OY	135	CAGGCGATCGCCCAAGAGAGATTGCCAGGAGCATCGCCAAATGGGGTGCGCCACAGAGCC	194	
Db	160634	CAGGCGATTCGCCAAGAAGAGACGCGGCCACAGGCGATTCGCCAAGAGAGACGCCGCCAGGCG	160575	
OY	195	ATCGCCATAGAGACGCCACCGAGGCGATTCGCCAACTGGAGCGCCGTCCAGGGCTTCGCC	254	
Db	160574	ATCGCCAACGAGAGACTTCGCCCCACGGAAATCGCACGAGAGACGCCGCCACGGCATTCGCC	160515	
OY	255	AACGGGGAAGCGCGTCTCAGCTTCGCCAACAGGGGAGCGCGGCCAGGGCATGCCAACGGG	314	
Db	160514	AGCCGAGAGCGCGCTCCAGAGGCGATCCCAAAGGAGGACGCGGCCAGGGCATGCCAACGAG	160455	
OY	315	GACGCGACCAAGGGCGATGGGCAACAGAGTTCACACATCCACGGCGATGCTTAACGAGAGCCG	374	
Db	160454	GACGCGCGCCAGGGGATTCGCCAAGGAGGACGCGGCCAGGGGATTCGCCAAGGAGAGCGCC	160395	
OY	375	GTCGAGGGCATTCGTAAACGAGGTGGCGCCGCCAGGGCATTCGCCAGAGAGACGCGGCCAG	434	
Db	160394	GCCCAAGGCGATTCGCCAAGAGGCGCCGCCGCCAGGGCATTCGCCAAGAAGAGCGCGCCAG	160335	
OY	435	GGAATTCGCCGAGGA----TGTGCGACAGGGGCAATCGCCACAAGAGAGCGCGCCAGGCGAT	490	
Db	160334	GGCATTCGCCCAACGAGAGATCGCCGCCAGGGCATTCGCCACAAGAGAGCGCGCCAGGAAAT	160275	
OY	491	GCGCAACGAG	500	
Db	160274	CGCCAGCGAG	160265	
RESULT	15			
LOCUS	AC008018	180884 bp	DNA	linear PRI 31-MAY-2000
DEFINITION	Homo sapiens chromosome 22q11 clone b379n11, complete sequence.			
ACCESSION	AC008018			
VERSION	AC008018.20	GI:7958973		
KEYWORDS	HTG:			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Euhayvota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mamula; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 180884)			
JOURNAL	Zhan,M. and Roe,B.A.			
REFERENCE	Homo sapiens Chromosome 22q11 BAC Clone b379n11 In BCRL2-GST Region			
AUTHORS	unpublished			
TITLE	2 (bases 1 to 180884)			
JOURNAL	Zhan,M. and Roe,B.A.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (10-JUL-1999) Department of Chemistry And Biochemistry,			
TITLE	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,			
JOURNAL	OK 73019, USA			
REFERENCE	3 (bases 1 to 180884)			
AUTHORS	Zhan,M. and Roe,B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-NOV-1999) Department Of Chemistry And Biochemistry,			
REFERENCE	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,			
AUTHORS	OK 73019, USA			
REFERENCE	4 (bases 1 to 180884)			
AUTHORS	Zhan,M. and Roe,B.A.			

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 715.867 Seconds
(without alignments)
14616.225 Million cell updates/sec

Title: US-09-489-101a-16_COPY_1_500
Perfect score: 500
Sequence: 1 gactctacgcatctctctgag.....cccgagggcatcgccacaaag 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mn:*
20: em_mn:*
21: em_mn:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgtg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	500	100.0	748	6 AX201595
2	498.4	99.7	2448	6 HSM801408
3	498.4	99.7	184590	2 AL356585
4	359.4	71.9	157086	2 AP000552
5	357.8	71.6	150754	9 AC023491
6	356.2	71.2	150724	9 AP000550
7	354.8	70.9	122364	9 AC007324
8	349.8	70.0	159550	2 AC013360
9	349.8	70.0	159550	2 AC013360
10	348.2	69.6	198557	2 AC008103
11	346.6	69.3	106850	9 AC007708
12	345	69.0	140876	9 AC009288
13	344	68.8	175158	9 AC007981
14	325.4	65.1	180884	9 AC008018
15	324.8	65.0	182617	9 AC007731
16	324.8	65.0	192592	9 AC005500
17	324.8	64.2	38849	9 AC024070
18	310	62.0	165050	9 AC007325
19	307.8	59.6	170102	9 AC008079
20	297.8	59.6	1647	9 HSM800994
21	297.8	59.6	1647	9 HSM801213
22	297.8	59.6	1647	9 HSM801213
23	265	53.0	28862	9 AC012398
24	265	53.0	176051	9 AC023490
25	259.6	51.9	143598	9 AC008132
26	250	50.0	94384	9 AC011718
27	229.6	45.9	175358	9 AC007981
28	180.2	35.0	202460	2 AC009066
29	177.6	35.5	160866	2 AC092532
30	171.6	34.3	69648	2 AC099825
31	131.2	26.2	110000	2 AL353694
32	116.4	23.3	390	6 AR135147
33	112.8	22.6	749	14 NP08P2
34	112.8	22.6	131995	14 OP075930
35	93.6	18.7	47852	1 MTW023
36	92.8	18.6	184590	2 AL356585
37	91.4	18.3	112509	2 AP003773
38	90	18.0	1647	9 HSM800994
39	90	18.0	1647	9 HSM801213
40	88.6	17.7	15348	1 HSM801408
41	87.4	17.5	15311	1 AE007163
42	87.4	17.5	15311	1 AE007164
43	87.2	17.4	1950	1 MTW023
44	87.2	17.4	1950	1 AB042821
45	87	17.4	122364	9 AC007324

ALIGNMENTS

RESULT	1	AX201595	748 bp	DNA	linear	PAT 30-NOV-2001
LOCUS	AX201595	Sequence 16 from Patent WO0153349.				
DEFINITION	AX201595					
ACCESSION	AX201595.1	GI:15391445				
VERSION						
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS		Stocker, E., Scallan, M. J., Jager, D., Old, L. J., Gure, A. O. and Chen, Y. T.				
TITLE		Small cell lung cancer associated antigens and uses therefor				
JOURNAL		Patent: WO 0153349-A 16 26-JUL-2001;				
		LUDWIG INSTITUTE FOR CANCER RESEARCH (US) ; MEMORIAL SLOAN-KETTERING CANCER CENTER (US) ; CORNELL RESEARCH FOUNDATION, INC. (US)				
FEATURES		Location/Qualifiers				
source		1..748				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				

BASE COUNT 166 a 261 c 246 g 72 t 3 others

ORIGIN

Query Match 100.0%; Score 500; DB 6; Length 748;
 Best Local Similarity 100.0%; Pred. No. 1,8e-36;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTACGCGATTGCTGAGACGCTGCCAGGCGATCGCTAATGAGACCGGACGAG 60
 |||||||
 DB 1 GAGCTACGCGATTGCTGAGACGCTGCCAGGCGATCGCTAATGAGACCGGACGAG 60
 |||||||

QY 61 GCATCGCTAATGAGACGACACCCAGTGCATCGCCACAGAGAGCCGCCAGGCGATCG 120
 |||||||
 DB 61 GCATCGCTAATGAGACGACACCCAGTGCATCGCCACAGAGAGCCGCCAGGCGATCG 120
 |||||||

QY 121 CCGAGAGCGCATTCAGGCGATCGCCACAGAGAGTGGCCAGGCGATCGCCATGAGG 180
 |||||||
 DB 121 CCGAGAGCGCATTCAGGCGATCGCCACAGAGAGTGGCCAGGCGATCGCCATGAGG 180
 |||||||

QY 181 TCGCGGACAGGCGATCGCCATGAGAGAGCGCCACAGGCGATCGCCATGAGGAG 240
 |||||||
 DB 181 TCGCGGACAGGCGATCGCCATGAGAGAGCGCCACAGGCGATCGCCATGAGGAG 240
 |||||||

QY 241 TCCAGGCGCTTCGCGACGCGGAGCGCGTCTCAGCTTCGCGACGCGGAGCGCG 300
 |||||||
 DB 241 TCCAGGCGCTTCGCGACGCGGAGCGCGTCTCAGCTTCGCGACGCGGAGCGCG 300
 |||||||

QY 301 GCATCGCGACAGGCGGAGCGCCACAGGCGATGAGGATCCAGGCGATCG 360
 |||||||
 DB 301 GCATCGCGACAGGCGGAGCGCCACAGGCGATGAGGATCCAGGCGATCG 360
 |||||||

QY 361 CTACAGGAGAGCGCGGAGCGGATCGCTAAGAGAGTGGCGCGCCAGGCGATCG 420
 |||||||
 DB 361 CTACAGGAGAGCGCGGAGCGGATCGCTAAGAGAGTGGCGCGCCAGGCGATCG 420
 |||||||

QY 421 AGAGCGCGCGCGGAGCGGATCGCGAGAGTGGCGAGATGCGCGAGGCGATCG 480
 |||||||
 DB 421 AGAGCGCGCGCGGAGCGGATCGCGAGAGTGGCGAGATGCGCGAGGCGATCG 480
 |||||||

QY 481 CCCAGGCGATCGCCACAGAG 500
 |||||||
 DB 481 CCCAGGCGATCGCCACAGAG 500
 |||||||

RESULT 2
 AX201594 1201 bp DNA linear PAT 30-AUG-2001
 LOCUS
 DEFINITION Sequence 15 from Patent WO0153349.
 ACCESSION AX201594
 VERSION AX201594.1 GI:15391443
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 STOCKER, E., SCANLAN, M.J., JAGER, D., OLD, L.J., GURE, A.O. and
 CHEN, Y.T.
 Small cell lung cancer associated antigens and uses therefor
 Patent: WO 0153349-A 15 26-JUL-2001.
 LUDWIG INSTITUTE FOR CANCER RESEARCH (US); MEMORIAL
 SLOAN-KETERING CANCER CENTER (US); CORNELL RESEARCH FOUNDATION,
 INC. (US)
 FEATURES
 source location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 255 a 377 c 349 g 178 t 42 others

BASE COUNT 255 a 377 c 349 g 178 t 42 others

ORIGIN

Query Match 100.0%; Score 500; DB 6; Length 1201;
 Best Local Similarity 100.0%; Pred. No. 1,6e-36;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTACGCGATTGCTGAGACGCTGCCAGGCGATCGCTAATGAGACCGGACGAG 60
 |||||||
 DB 1 GAGCTACGCGATTGCTGAGACGCTGCCAGGCGATCGCTAATGAGACCGGACGAG 60
 |||||||

QY 61 GCATCGCTAATGAGAGACACCCAGTGCATCGCCACAGAGAGCCGCCAGGCGATCG 120
 |||||||
 DB 61 GCATCGCTAATGAGAGACACCCAGTGCATCGCCACAGAGAGCCGCCAGGCGATCG 120
 |||||||

QY 121 CCGAGAGCGCATTCAGGCGATCGCCACAGAGAGTGGCCAGGCGATCGCCATGAGG 180
 |||||||
 DB 121 CCGAGAGCGCATTCAGGCGATCGCCACAGAGAGTGGCCAGGCGATCGCCATGAGG 180
 |||||||

QY 181 TCGCGGACAGGCGATCGCCATGAGAGAGCGCCACAGGCGATCGCCATGAGGAG 240
 |||||||
 DB 181 TCGCGGACAGGCGATCGCCATGAGAGAGCGCCACAGGCGATCGCCATGAGGAG 240
 |||||||

QY 241 TCCAGGCGCTTCGCGACGCGGAGCGCGTCTCAGCTTCGCGACGCGGAGCGCG 300
 |||||||
 DB 241 TCCAGGCGCTTCGCGACGCGGAGCGCGTCTCAGCTTCGCGACGCGGAGCGCG 300
 |||||||

QY 301 GCATCGCGACAGGCGGAGCGCCACAGGCGATGAGGATCCAGGCGATCG 360
 |||||||
 DB 301 GCATCGCGACAGGCGGAGCGCCACAGGCGATGAGGATCCAGGCGATCG 360
 |||||||

QY 361 CTACAGGAGAGCGCGGAGCGGATCGCTAAGAGAGTGGCGCGCCAGGCGATCG 420
 |||||||
 DB 361 CTACAGGAGAGCGCGGAGCGGATCGCTAAGAGAGTGGCGCGCCAGGCGATCG 420
 |||||||

QY 421 AGAGCGCGCGCGGAGCGGATCGCGAGAGTGGCGAGATGCGCGAGGCGATCG 480
 |||||||
 DB 421 AGAGCGCGCGCGGAGCGGATCGCGAGAGTGGCGAGATGCGCGAGGCGATCG 480
 |||||||

QY 481 CCCAGGCGATCGCCACAGAG 500
 |||||||
 DB 481 CCCAGGCGATCGCCACAGAG 500
 |||||||

RESULT 3
 HSM801408 2448 bp mRNA linear PRI 18-FEB-2000
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKF2p434C196 (from clone DKF2p434C196);
 partial cds.
 ACCESSION AL133561
 VERSION AL133561.1 GI:6599133
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2448)
 KOEHLER, K., BEYER, A., MEWES, H.W., GASENHUBER, J. and WIEMANN, S.
 Direct Submission
 Submitted (15-DEC-1999) MIPS, Am Klopfersplitz 18a, D-82152
 Martinsried, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 This clone (DKFZ434C196) is available at the RZPD in Berlin.
 Please contact the RZPD, Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cdmv/>.
 FEATURES
 source location/Qualifiers
 1..2448
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKF2p434C196"
 /issue="1"
 /release="1999"
 /clone_lib="434 (synonym: htes3). Vector pSPORT1; host
 DH10B; sites NotI + SalI"

repeat_region	/evidence-not_experimental 11054. .11343 /rpt_family="AlusX"	repeat_region	complement(20885. .21149) /rpt_family="AluY"
repeat_region	/evidence-not_experimental 11344. .11367 /rpt_family="TTAA)n"	repeat_region	/evidence-not_experimental 21186. .21481 /rpt_family="AluY"
repeat_region	/evidence-not_experimental complement(12418. .12736) /rpt_family="AlusG"	repeat_region	/evidence-not_experimental complement(21482. .21758) /rpt_family="AlusG"
repeat_region	/evidence-not_experimental complement(13993. .14328) /rpt_family="AluJo"	repeat_region	/evidence-not_experimental complement(21759. .21862) /rpt_family="AlusP/q"
repeat_region	/evidence-not_experimental complement(14759. .14875) /rpt_family="FLAM.C"	repeat_region	/evidence-not_experimental 24439. .24588 /rpt_family="AluY"
repeat_region	/evidence-not_experimental 14876. .14886 /rpt_family="Charlie4"	repeat_region	/evidence-not_experimental complement(24678. .24968) /rpt_family="AlusX"
repeat_region	/evidence-not_experimental complement(14687. .15155) /rpt_family="AlusX"	repeat_region	/evidence-not_experimental complement(25016. .25306) /rpt_family="AluYb8"
repeat_region	/evidence-not_experimental 15156. .15239 /rpt_family="Charlie4"	repeat_region	/evidence-not_experimental complement(25310. .25609) /rpt_family="AluYb8"
repeat_region	/evidence-not_experimental 15618. .15788 /rpt_family="L2"	repeat_region	/evidence-not_experimental 25627. .25853 /rpt_family="AlusX"
repeat_region	/evidence-not_experimental 15911. .15924 /rpt_family="L1M87"	repeat_region	/evidence-not_experimental 25858. .25885 /rpt_family="(A)n"
repeat_region	/evidence-not_experimental complement(15925. .16230) /rpt_family="AluJo"	repeat_region	/evidence-not_experimental 26308. .26588 /rpt_family="AluJo"
repeat_region	/evidence-not_experimental 16231. .16940 /rpt_family="L1M87"	repeat_region	/evidence-not_experimental complement(30748. .31041) /rpt_family="AluJo"
repeat_region	/evidence-not_experimental complement(17656. .17816) /rpt_family="FLAM.A"	repeat_region	/evidence-not_experimental 35559. .35661 /rpt_family="AlusG/x"
repeat_region	/evidence-not_experimental complement(17861. .18018) /rpt_family="AluJo"	repeat_region	/evidence-not_experimental complement(35833. .35990) /rpt_family="MER5B"
repeat_region	/evidence-not_experimental 18019. .18049 /rpt_family="(TTTAA)n"	repeat_region	/evidence-not_experimental 36421. .36721 /rpt_family="AlusG"
repeat_region	/evidence-not_experimental complement(18050. .18176) /rpt_family="AluJo"	repeat_region	/evidence-not_experimental 37016. .37083 /rpt_family="MUT1E"
repeat_region	/evidence-not_experimental 18300. .18602 /rpt_family="AluJo"	repeat_region	/evidence-not_experimental 37175. .37372 /rpt_family="MUT1E2"
repeat_region	/evidence-not_experimental complement(19466. .19612) /rpt_family="L2"	repeat_region	/evidence-not_experimental 37480. .37793 /rpt_family="AluY"
repeat_region	/evidence-not_experimental 19721. .20022 /rpt_family="AluY"	repeat_region	/evidence-not_experimental 38069. .38351 /rpt_family="AlusG"
repeat_region	/evidence-not_experimental complement(20025. .20302) /rpt_family="AlusG"	repeat_region	/evidence-not_experimental 38506. .38612 /rpt_family="MIR"
repeat_region	/evidence-not_experimental complement(20308. .20325) /rpt_family="AlusP/q"	repeat_region	/evidence-not_experimental 38841. .38985 /rpt_family="AlusG/x"
repeat_region	/evidence-not_experimental 20326. .20374 /rpt_family="(TG)n"	repeat_region	/evidence-not_experimental 38987. .39118 /rpt_family="L1PA16"
repeat_region	/evidence-not_experimental complement(20375. .20445) /rpt_family="AlusP/q"	repeat_region	/evidence-not_experimental 39119. .39173 /rpt_family="L1PA16"
repeat_region	/evidence-not_experimental		

Query Match 71.9%; Score 359.4; DB 9; Length 157086;
 Best Local Similarity 86.3%; Pred. No. 1.1e-24;
 Matches 422; Conservative 0; Mismatches 61; Indels 6; Gaps 2;

	RESULT	b		
	LOCUS	AC023491/c		
	DEFINITION	Homo sapiens chromosome 22q11 clone unknown, complete sequence.		
	VERSION	AC023491		
	KEYWORDS	AC023491.21 GI:9309521		
	SOURCE	HTG.		
	ORGANISM	human.		
	REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	TITLE	1 (bases 1 to 150754)		
	JOURNAL	Hu,P., Wu,H., Yang,L., Morrow,B.E. and Roe,B.A. Homo sapiens Chromosome 22q11 BAC Clone 659mII In BCRI2-GGT Region Unpublished		
	AUTHORS	2 (bases 1 to 150754) Hu,P., Wu,H., Yang,L., Morrow,B.E. and Roe,B.A.		
	TITLE	Direct Submission		
	JOURNAL	Submitted (14-FEB-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
	REFERENCE	3 (bases 1 to 150754) Hu,P., Wu,H., Yang,L., Morrow,B.E. and Roe,B.A.		
	AUTHORS	Direct Submission		
	TITLE	Submitted (26-JUN-2000) Department Of Chemistry And Biochemistry,		
	JOURNAL	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
	REFERENCE	4 (bases 1 to 150754) Hu,P., Wu,H., Yang,L., Morrow,B.E. and Roe,B.A.		
	AUTHORS	Direct Submission		
	TITLE	Submitted (01-JUL-2000) Department Of Chemistry And Biochemistry,		
	JOURNAL	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
	REFERENCE	5 (bases 1 to 150754) Hu,P., Wu,H., Yang,L., Morrow,B.E. and Roe,B.A.		
	AUTHORS	Direct Submission		
	TITLE	Submitted (01-JUL-2000) Department Of Chemistry And Biochemistry,		
	JOURNAL	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA		
	REFERENCE	Duplicate Submission		
	AUTHORS	Hu,P., Wu,H., Yang,L., Morrow,B.E. and Roe,B.A.		
	TITLE	Direct Submission		

	Query Match	71.6%	Score 357.8	DB 9	Length 150754
	Best Local Similarity	86.1%	Pred. No. 1.5e-24		
	Matches 421	Conservative	0	Mismatches 62	Indels 6
QY 18	GAGGAGCCTCCCGAGGCGATCGCTATGAGGACGCCGACAGGGCATTCGTTAATGAGGAC	77			
Db 11233	GAGGAGCGCCGCCGAGGCGATTCGCCAACGAGGACGCCGCCCGAGGGCATTCGCCAAGAGAGAC	111174			
QY 78	ACACACCGATGCTCGCCACGAGGAAACCGCCCGAGGGCATTCGCC---GAGAGCGCATTC	134			
Db 111173	GCCGCCGAGGCGATTCGCCAAGAGAGAGCGCCGCCCGAGGGCATTCGCCAAGAGAGAGCGCCGC	111114			
QY 135	CAGGCGCATTCGCCAAGAGAGAGTTGCCCGAGGGCATTCGCCAATGGGGTTCGCCGCAAGAGC	194			
Db 111113	CAGGCGATTCGCCAAGAGAGAGCGGCCGCCACGGAAATTCGCAACGAGAGAGCGCCGCCAAGG	111054			
QY 195	ATCGCCATATGAGGAGCGCACCCAGGGCATTCGCGCAATCGGACGCCGCTCCAGCGGCTTCGCC	254			
Db 111053	ATCGCCAAAGCAGAGACGGCCCCCGAGGGCATTCGCCACGAGAGAGCGGCCGCCCGAGGGCATTCGAC	110994			
QY 255	AACGGGAGCGCGCTCTCAGCTTCGCCCAACGGGGGAGCGGCCGCCCGAGGGCATTCGCCCAACGGG	314			
Db 110993	AAGGAGAGCGCCGCCCGAGGGCATTCGCCCAACGAGGAGCGGCCGCCCGAGGGCATTCGCCCAACGGAG	110934			
QY 315	GAGGCCACCAAGGGCATTCGGGCAACGAGGATCAATCCACGGCATTCGGTTACGAGAGAGAGCGC	374			
Db 110933	GAGCGCGCCACGAGGCGATTCGCCAAGAGAGAGCGCATTCACAGGGCATTCGCCAAGAGATACGCC	110874			
QY 375	GTCGAGGGGATCGCTAAGAGGTGGCGCGCCCGAGGGCATTCGCCCAACGAGAGAGCGCCGCCAG	434			
Db 110873	GTCGAGGCGATTCGCCAAGAGAGAGCGGCCGCCCGAGGGCATTCGCCAAGAGAGAGCGCGCCCGAG	110814			
QY 435	GGAATCGCC---GAGGATTCGCAACAGGGCATTCGCCCAACGAGAGAGCGCGCCCGAGGGCATTC	491			
Db 110813	GGGATCGCCCAAGAGAGAGACCGCGCCCGAGGGCATTCGCCAAGAGAGAGAGCGCGCCCGAGGGCATTC	110754			
QY 492	GCCACACAG 500				
Db 110753	GCCACACAG 110745				
RESULT 7	AP000550	150724 bp	DNA	linear	PRI 01-OCT-1999
LOCUS	AP000550	150724 bp	DNA	linear	PRI 01-OCT-1999
DEFINITION	Homio sapiens genomic DNA, chromosome 22q11.2, BCR12 region,				
	clone:KH1592A4.				
ACCESSION	AP000550				
VERSION	AP000550.1	GI:5931536			
KEYWORDS	homio sapiens pre-pro-B cell cell_line:FLB14-14 DNA, clone_lib:Keio				
SOURCE	BAC library clone:KH1592A4.				
ORGANISM	Homio sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

```

REFERENCE 1 (bases 1 to 150724)
AUTHORS Shimizu,N.
TITLE Human DNA sequence from clone KB1592A4 on chromosome 22q11.2
JOURNAL Published Only in Database (1999) In press
REFERENCE 2 (bases 1 to 150724)
AUTHORS Shimizu,N.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1999) to the DDBJ/EMBL/GenBank databases.
Nobuyoshi Shimizu, Keio University, School of Medicine, Molecular
Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-0016, Japan
Fax:81-3-3351-2370
(E-mail:nshimizu@med.keio.ac.jp, Tel:81-3-3351-2370,
This is a complete sequence of the insert of KB1592A4 clone. The
distal adjacent clone is KB876E2 (Acc.#AF00551) with 1196-bp
overlapping.

COMMENT
FEATURES
    source
        1..150724
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="FEB14-14"
        /cell_type="pre-pro-B cell"
        /chromosome="22"
        /clone="KB1592A4"
        /map="22q11.2"
        /clone_id="Keio BAC library"
        /complement(1..90)
        /rpt_family="Alus"
        /evidence-not_experimental
        /complement(99..399)
        /rpt_family="Alusg"
        /evidence-not_experimental
        /complement(401..570)
        /rpt_family="L1R33"
        /evidence-not_experimental
        /complement(572..640)
        /rpt_family="MIR"
        /evidence-not_experimental
        /complement(649..716)
        /rpt_family="(TCCA)n"
        /evidence-not_experimental
        /complement(717..752)
        /rpt_family="L2"
        /evidence-not_experimental
        /complement(757..862)
        /rpt_family="MIR"
        /evidence-not_experimental
        /complement(933..1059)
        /rpt_family="L2"
        /evidence-not_experimental
        /complement(1955..2151)
        /rpt_family="AluNb"
        /evidence-not_experimental
        /complement(2578..2875)
        /rpt_family="Alusx"
        /evidence-not_experimental
        /complement(3310..3408)
        /rpt_family="L1R16A"
        /evidence-not_experimental
        /complement(4424..4591)
        /rpt_family="L1PA4"
        /evidence-not_experimental
        /complement(4617..4637)
        /rpt_family="(TAAA)n"
        /evidence-not_experimental
        /complement(4663..5286)
        /rpt_family="L1"
        /evidence-not_experimental
        /complement(5353..5554)
        /rpt_family="L1"
        /evidence-not_experimental
        /complement(5542..5625)
        /rpt_family="L1M1"

repeat_region /evidence-not_experimental
complement(5626..5882)
/rpt_family="AluNb"
/evidence-not_experimental
complement(5883..6008)
/rpt_family="L1M1"
/evidence-not_experimental
complement(6025..6235)
/rpt_family="AluNb"
/evidence-not_experimental
complement(6236..6447)
/rpt_family="L1M2"
/evidence-not_experimental
complement(6486..6554)
/rpt_family="Alu/FRAM"
/evidence-not_experimental
complement(6556..6709)
/rpt_family="L1"
/evidence-not_experimental
complement(6778..7721)
/rpt_family="L1M6"
/evidence-not_experimental
complement(8075..8296)
/rpt_family="L1PA8"
/evidence-not_experimental
complement(8679..8980)
/rpt_family="Alusx"
/evidence-not_experimental
complement(9224..9298)
/rpt_family="L1M2"
/evidence-not_experimental
complement(9324..9932)
/rpt_family="L1M2"
/evidence-not_experimental
complement(10013..10508)
/rpt_family="L1M2"
/evidence-not_experimental
complement(10548..10734)
/rpt_family="MIR"
/evidence-not_experimental
complement(10867..11330)
/rpt_family="L1MA7"
/evidence-not_experimental
complement(11332..11537)
/rpt_family="L1M2"
/evidence-not_experimental
complement(11540..11728)
/rpt_family="Alusg/x"
/evidence-not_experimental
complement(11730..11910)
/rpt_family="L1M2"
/evidence-not_experimental
complement(11911..12230)
/rpt_family="L1PA6"
/evidence-not_experimental
complement(12231..12416)
/rpt_family="L1P"
/evidence-not_experimental
complement(12640..12746)
/rpt_family="MIR"
/evidence-not_experimental
complement(12901..13183)
/rpt_family="Alusg"
/evidence-not_experimental
complement(13459..13774)
/rpt_family="Aluv"
/evidence-not_experimental
complement(13882..14031)
/rpt_family="MTR1E2"
/evidence-not_experimental
complement(14171..14238)
/rpt_family="MTR1B"
/evidence-not_experimental

```

```

repeat_region complement(14533, .14830)
                /rpt_family="AluSg"
                /evidence=not_experimental
                15255, .15418
                /rpt_family="MER5B"
repeat_region /evidence=not_experimental
                complement(15590, .15692)
                /rpt_family="AluSg/X"
                /evidence=not_experimental
                20210, .20503
                /rpt_family="AluSg"
repeat_region /evidence=not_experimental
                complement(24663, .24943)
                /rpt_family="AluSg"
                /evidence=not_experimental
                25366, .25393
                /rpt_family="(?)n"
                /evidence=not_experimental
                complement(25398, .25624)
                /rpt_family="AluSx"
                /evidence=not_experimental
                25642, .25941
                /rpt_family="AluSx"
repeat_region /evidence=not_experimental
                complement(25945, .26235)
                /rpt_family="AluYb8"
                /evidence=not_experimental
                26283, .26574
                /rpt_family="AluSx"
                /evidence=not_experimental
                complement(26564, .26814)
                /rpt_family="AluY"
                /evidence=not_experimental
                26991, .29494
                /rpt_family="AluSg/q"
                /evidence=not_experimental
                29495, .29772
                /rpt_family="AluSg"
                /evidence=not_experimental
                complement(29773, .30076)
                /rpt_family="AluY"
                /evidence=not_experimental
                30113, .30372
                /rpt_family="AluY"
repeat_region

Query Match 71.2%; Score 356.2; DB 9; Length 150724;
Best Local Similarity 85.9%; Pred. No. 2.1e-24;
Matches 420; Conservative 0; Mismatches 63; Indels 6; Gaps 2;

OY 18 GAGGAGCGTCCGAGGATCGCTATGAGGAGCGGACGAGGATCGCTATGAGGAG 77
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84631 GAGGAGCGTCCGAGGATCGCTATGAGGAGCGGACGAGGATCGCTATGAGGAG 84690

OY 78 ACCACCCAGTGCATCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 134
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84691 GCCGCCAGGAGGATCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84750

OY 135 CAGGAGCGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 194
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84751 CAGGAGCGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84810

OY 195 ATGCCCATAGAGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 254
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84811 ATGCCCATAGAGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84870

OY 255 AACGGGAGCGCTCTCTCAGCTTCGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAG 314
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84871 AGCGAGGAGCGCTCTCTCAGCTTCGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAG 84930

OY 315 GAGCGCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 374
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84931 GAGCGCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84990

```

```

OY 375 GTCCAGGAGGATCGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 434
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84991 GCCCAGGAGGATCGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 85050

OY 435 GGAGTGGCC---GAGATGTCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 491
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 85051 GGAGTGGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 85110

OY 492 GCCAACAG 500
    ||||||| ||
Db 85111 GCCAACAG 85119

RESULT 8
LOCUS AC007324 122364 bp DNA linear PRI 18-MAY-2000
DEFINITION Homo sapiens chromosome 22q11 clone b293, complete sequence.
ACCESSION AC007324
VERSION AC007324.55 GI:7923342
KEYWORDS HYG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 122364)
Fu,Y., Fang,F., Wang,Q., Pan,H., McDermid,H. and Roe,B.A.
Submitted (19-DEC-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 122364)
Fu,Y., Fang,F., Wang,Q., Pan,H., McDermid,H. and Roe,B.A.
Submitted (06-MAY-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 122364)
Fu,Y., Fang,F., Wang,Q., Pan,H., McDermid,H. and Roe,B.A.
Submitted (18-MAY-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 18, 2000 this sequence version replaced gi:7712128.

COMMENT
FEATURES
    source
        1..122364
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="22q11"
        /clone="b293"

BASE COUNT 31795 a 28253 c 28482 g 33834 t

Query Match 70.9%; Score 354.6; DB 9; Length 122364;
Best Local Similarity 85.7%; Pred. No. 3.1e-24;
Matches 419; Conservative 0; Mismatches 64; Indels 6; Gaps 2;

OY 18 GAGGAGCGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 77
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 82628 GAGGAGCGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 82687

OY 78 ACCACCCAGTGCATCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 134
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 82688 GCCGCCAGGAGGATCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 82747

```

```

QY 135 CAGGCGATCGCCACGAGAGGTTGCCAGGGCATCGCCATGGGTCGCCGACGAGGC 194
DB 83748 CAGGCGATCGCCACGAGAGGAGGCGCCGCGGCGATCGCCACGAGAGGCGCCGCGACGAGGC 82807
QY 195 ATGCGCAATGAGGAGCGCCAGGCGCATCGCCACACTGGAGCGCGTTCACAGCGCTTCCGC 254
DB 82808 ATGCGCAATGAGGAGCGCGCGCGCGCGCATCGCCACGAGAGGAGGCGCGCGCGCGCGCGCG 82867
QY 255 AACGGGAGCGCGCGTCTCAGCTTCGCGCAAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314
DB 82868 AACGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 82927
QY 315 GACGCGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 374
DB 82928 GACGCGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 82987
QY 375 GTCCAGGCGCATCGCTTACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 434
DB 82988 GTCCAGGCGCATCGCTTACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 83047
QY 435 GGAATCGCC--GAGGATGTCGCGCAGAGGCGCATCGCCACGAGAGGCGCGCGCGCGCGCGCATC 491
DB 83048 GGAATCGCC--GAGGATGTCGCGCAGAGGCGCATCGCCACGAGAGGCGCGCGCGCGCGCGCATC 83107
QY 492 GCCACACAG 500
DB 83108 GCCACACAG 83116

RESULT 9
AC013360 159550 bp DNA linear HTG 04-MAY-2001
LOCUS AC013360 Homo sapiens chromosome 11, clone RP11-278E23 map 11, WORKING DRAFT
DEFINITION AC013360.13 unordered pieces.
ACCESSION AC013360.4 GI:7229982
VERSION AC013360.4
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159550)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 11, clone RP11-278E23
REFERENCE 2 (bases 1 to 159550)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguski,M., Bouhassira,D.,
Brown,A., Castle,A., Collins,S., Collins,S., Collins,S., Collins,S.,
Cooke,P., Dearlano,K., Desai,K., Domini,M., Donnelly,L., Doyle,M.,
Ferreira,P., Fitzhugh,M., Forrest,C., Funke,R., Gage,D., Gage,D.,
Galligan,J., Gargano,S., Grant,G., Hagos,B., Heath,D., Horton,L.,
Howard,J., Johnson,R., Jones,C., Kahn,L., Karas,A., Klein,D.,
Lehoczky,J., Liou,C., Locke,K., Macdonald,P., Margulis,N.,
McMurry,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessier,S., Tjelle,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:6479088.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www.sag.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information
Center project name: L2604
Center clone name: 278_E_23
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960721
Consensus quality: 144663 bases at least Q40
Consensus quality: 153318 bases at least Q30
Consensus quality: 156484 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 158350; sum-of-contigs
Quality coverage: 5.4 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1729: contig of 1729 bp in length
1730 1829: gap of 100 bp
1830 3716: contig of 1887 bp in length
3717 3816: gap of 100 bp
3817 6950: contig of 3134 bp in length
6951 7051: gap of 100 bp
7051 9136: contig of 2086 bp in length
9137 9236: gap of 100 bp
9237 17859: contig of 8623 bp in length
17860 17959: gap of 100 bp
17960 326939: contig of 8980 bp in length
326940 27039: gap of 100 bp
27040 36269: contig of 9230 bp in length
36270 36369: gap of 100 bp
36370 48574: contig of 12205 bp in length
48575 48674: gap of 100 bp
48675 62315: contig of 13641 bp in length
62316 62415: gap of 100 bp
62416 77457: contig of 15042 bp in length
77458 77557: gap of 100 bp
77558 96884: contig of 19327 bp in length
96885 96984: gap of 100 bp
96985 124230: contig of 27246 bp in length
124231 124330: gap of 100 bp
124331 159550: contig of 35220 bp in length.
-----
FEATURES
source
1. 159550
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-278E23"
/clone_id="RP11-11 Human Male BAC"
1. 1729
/note="assembly-fragment"
1830. 3716
/note="assembly-fragment"
3817. 6950
/note="assembly-fragment"
7051. 9136
/note="assembly-fragment"
9237. 17859
/note="assembly-fragment"
17960. 326939
/note="assembly-fragment"
27040. 36269
/note="assembly-fragment"
36370. 48574
/note="assembly-fragment"
48675. 62315

```


REFERENCE
AUTHORS
TITLE
JOURNAL
OK 73019, USA
10 (bases 1 to 98557)
Wang, Q., Fang, F., Fu, Y., Pan, H., Edelman, L. and Roe, B.A.
Submitted (05-FEB-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
11 (bases 1 to 98557)
Wang, Q., Fang, F., Fu, Y., Pan, H., Edelman, L. and Roe, B.A.
Submitted (08-FEB-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
12 (bases 1 to 98557)
Wang, Q., Fang, F., Fu, Y., Pan, H., Edelman, L. and Roe, B.A.
Submitted (10-FEB-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
13 (bases 1 to 98557)
Wang, Q., Fang, F., Fu, Y., Pan, H., Edelman, L. and Roe, B.A.
Submitted (14-FEB-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
14 (bases 1 to 98557)
Wang, Q., Fang, F., Fu, Y., Pan, H., Edelman, L. and Roe, B.A.
Submitted (12-MAR-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
15 (bases 1 to 98557)
Wang, Q., Fang, F., Fu, Y., Pan, H., Edelman, L. and Roe, B.A.
Submitted (27-MAY-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT
On Mar 12, 2000 this sequence version replaced gi:6938024.
Because these overlapping clones came from different libraries
there are numerous instances of insertions, deletions, and single
nucleotide polymorphisms in the overlapping regions below.
AC008133(pac69506) 136107 145670 (0) overlaps AC008103(pac69911) 1
9556 (889001) AC007981(b291) 132871 175358 (0) overlaps
AC008103(pac69911) 1 44355 (54202) AC007325(b48) 71505 165050 (0)
overlaps AC008103(pac69911) 1 94144 (4413) AC008103(pac69911) 56194
98557 (0) overlaps AC007326(p423) 1 42365 (60533).
Location/Qualifiers
1. .98557
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="22q12"
/clone="pac69911"

BASE COUNT 25600 a 24617 c 24180 g 24160 t

ORIGIN

Query Match 69.6%; Score 348.2; DB 9; Length 98557;
Best Local Similarity 84.9%; Pred. No. 1,2e-23;
Matches 415; Conservative 0; Mismatches 68; Indels 6; Gaps 2;

QY 18 GAGAGCGCTCCGAGGCGATCGCTAATAGAGAGCGGAGGCGATCGCTAATAGAGAC 77
DB 27916 GAGAGCGCTCCGAGGCGATCGCTAATAGAGAGCGGAGGCGATCGCTAATAGAGAC 27857

QY 78 ACACACCATGATCGGCGAGGAGAGAGCGGCGGAGGCGATCGCTAATAGAGAC 134
DB 27856 GCGCGCGAGGCGGAGGAGAGAGCGGCGGAGGCGATCGCTAATAGAGAGCGGCGG 27797

QY 135 CAGGCGATCGGCGAGGAGGAGGAGGCGGAGGCGATCGGCGGAGGCGAGGAGG 194
DB 27796 CAGGCGATCGGCGAGGAGGAGGAGGCGGAGGCGATCGGCGGAGGCGAGGAGG 27737

QY 195 ATCCGCAATAGAGAGCGGCGGAGGCGATCGGCGGAGGCGGAGGCGGAGGCGG 254
||||||| ||||||| ||| ||||||| ||| ||| ||||||| |||||

DB 27736 ATCCGCAATAGAGAGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 27617

QY 255 AAGGGAGGCGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 314
||| ||||||| ||| ||| ||||||| ||| ||||||| ||||||| ||| |||

DB 27616 AGGAGAGCGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 27617

QY 315 GAGCGCGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 374
||| ||||||| ||| ||| ||||||| ||| ||||||| ||||||| ||| |||

DB 27616 GAGCGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 27557

QY 375 GTCAGGCGGATCGCTAGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 434
||||| ||||||| ||| ||| ||||||| ||| ||||||| ||||||| ||| |||

DB 27556 GCGGAGGCGGATCGCTAGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 27497

QY 435 GGATCGCGC---GAGCATGTGCGACAGGCGGATCGGCGGAGGCGGAGGCGGAGGCGGATC 491
||| ||||||| ||| ||| ||||||| ||| ||||||| ||||||| ||| |||

DB 27496 GCGATCGCGCAAGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGATC 27437

QY 492 GCCAACAAG 500
||||| ||

DB 27436 GCCAACAAG 27428

RESULT 12
AC007708/c 106650 bp DNA linear PRI 31-MAY-2000
LOCUS
DEFINITION Homo sapiens chromosome 22q11 clone b563b9, complete sequence.
AC007708
VERSION
KEYWORDS
HNG.
SOURCE
human.
ORANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
1 (bases 1 to 106650)
Zhan, M. and Roe, B.A.
Unpublished
2 (bases 1 to 106650)
Zhan, M. and Roe, B.A.
Direct Submission
Submitted (02-JUN-1999) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
3 (bases 1 to 106650)
Zhan, M. and Roe, B.A.
Direct Submission
Submitted (25-SEP-1999) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 106650)
Zhan, M. and Roe, B.A.
Direct Submission
Submitted (13-MAY-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 106650)
Zhan, M. and Roe, B.A.
Direct Submission
Submitted (31-MAY-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Sep 25, 1999 this sequence version replaced gi:5919292.
Because these overlapping clones came from different libraries
there are numerous instances of insertions, deletions, and single
nucleotide polymorphisms in the overlapping regions below.
AC008018(b375n11) 146090 180888 (0) overlaps AC007708(b563b9) 1
34927 (71723) AC007708(b563b9) 1843 106650 (0) overlaps
AC009288(p413m7) 1 101851 (39025).
Location/Qualifiers
1. .106650
/organism="Homo sapiens"
/db_xref="taxon:9606"

Fri Oct 11 09:30:46 2002

us-09-489-101a-16_copy_1_500.rge

Page 15

TITLE	Direct Submission
JOURNAL	Submitted (19-MAY

Direct Submission
Submitted (19-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

5 (bases 1 to 180884)
Zhan, M. and Roe, B.A.
Direct Submission
Submitted (20-MAY-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

6 (bases 1 to 180884)
Zhan, M. and Roe, B.A.
Direct Submission
Submitted (31-MAY-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT

On May 20/0000 this sequence version replaced g1:7940354. Because these overlapping clones came from different libraries there are numerous instances of insertions, deletions, and single nucleotide polymorphisms in the overlapping regions below.

AC002472(P-n5) 122927 147086 (0) overlaps AC008018(b379n11) 1 241661
(156723) AC007664(b45314) 79011 162470 (0) overlaps
(AC008018(b379n11) 1 83457 (97431) AC008018(b379n11) 6082 49563
(130925) overlaps AC002054(clone88) 1 43881 (0) AC008018(b379n11)
13707 22565 (158325) AC002049(10747) 10173 19026 (0)
AC008018(b379n11) 20528 54785 (126103) overlaps AC002051(26a1) 1
34538 (0) AC008018(b379n11) 146090 180888 (0) overlaps
AC007708(b56389) 1 349527 (71123) AC008018(b379n11) 147946 180888
(0) overlaps AC009288(f413m7) 1 32943 (107933).

FEATURES
SOURCE

```
/organism="Homo sapiens";  
/db_xref="taxon:9606"
```

BASE COUNT	48190 a	42579 c	41490 g	48625 t
ORIGIN				

Query Match	65.18;	Score 325.4;	DB 9;	Length 180884;
Best Local Similarity	81.0%;	Pred. No. 1.1e-21;		
Matches 404;	Conservative 0;	Mismatches 91;	Indels 4;	Gaps 2;

Oy 5 CTACGGCATTGCTAGAGACCGCTGCCCAAGGGCATCGCTATATGAGACGCCACCCAGGGCAT 64
| | | | |
Db 103093 CGACATCGCTACAGSAGATACTGTACACGACATTCGCTAACGAGSCCGGCACAAGSGCAT 103157

QY 65 CGTATGAGGACACCACCCAGTGCATCGCCACGAGGAAGCCGCCCAGGGCATCGC - - 121

Db 103153 CGCCAACGAGGATGCCCGGCCCAAGGCAATCGCCACGAGGACGCGCCCAAGGGCATCGCCAG 1032112

Db 103213 CGAGGACGCCGCCACGGCATGCCAGGAGGACGCCGCCACGGCATGCCAGGAGGA 103272

Oy 182 CGCCGCACAGGGCATTGCCAATGAGACGCCACCAGGGCATGCCCAACTGGGAGCCCGT 241
||||| | ||||| | | ||||| | ||||| | ||||| | |||||
Db 103273 GCGCGCTCACGGCATGTGGCAGGAGAAAGCCGCTCAGGGCATGGCAACAGAGGAGCCCGT 10333

017 2A2 CCACCCTTCCACAAACGCGACGCGCAGCTCACATTCAGCAACGCGCAGCG 301

```

0x    303  CATTGCCGATACCGCCAGCAGCAAGGCCTCACGGCATTCGCGAACGAGACGCGCCAGGG  361
Db 10333 CCAGGCGATCGCCCAAGGAGGAGCGCCGTCCACGCGCATTCGCGAACGAGACGCGCCAGGG  10339

```

06 103303 ***** 103457

QY 362 TACGAGACCGCTCCAGGCATCGCTAACGAGGTGGCCGCCACGAGGCATCGCCACGA 421

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 103513 GGACGGCCGACGGCATCGCCCAACGA-GCTGTATACGACATCGTAACGAGGACACTT 10357

QY 482 CCAGGGCATGCCAACAAG 500
 ||| | |||| |||||
 Db 103572 ACAAGCCGTCGCTACAAG 103590

Search completed: October 10, 2002, 18:51:52
Job time : 1325.87 secs

Fri Oct 11 09:30:52 2002

us-09-489-101a-17_copy_1_500.rni

Page 2

```

RESULT 2
US-08-659-188-17/c
; Sequence 17, Application US/08659188
; Patent No. 6002069
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; TITLE OF INVENTION: Seed Plants Exhibiting Inducible Early
; TITLE OF INVENTION: Reproductive Development and Methods of Making Same
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,188
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1946
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1651
; NAME/KEY: misc_feature
; LOCATION: 1..1656
; OTHER INFORMATION: /note= "domain = ecdysone receptor
; OTHER INFORMATION: ligand binding domain."
US-08-659-188-17

Query Match          20.8%; Score 104.2; DB 3; Length 1656;
Best Local Similarity 67.8%; Pred. No. 8.2e-14;
Matches 145; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 10 GGGCTGGGGCTGGGGCTGGAGCTGAGCTGAGCTGAGGCTGGGGCTGGGGCTGG 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1341 GGGCAGCGAGCGAGAGCGAGAGCGAGCTGTGCTGTGCTGTGCTGTGCTGTG 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 GGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1281 CGTCTGAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 AGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1221 AGGTTGTAGCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 GGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1161 GGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-08-655-227-17/c

```

```

; Sequence 17, Application US/08655227
; Patent No. 6025483
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; TITLE OF INVENTION: Maize and Cauliflower APEPTAL Gene
; TITLE OF INVENTION: Products and Nucleic Acid Molecules Encoding Same
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/655,227
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2143
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1651
; NAME/KEY: misc_feature
; LOCATION: 1..1656
; OTHER INFORMATION: /note= "domain = ecdysone receptor
; OTHER INFORMATION: ligand binding domain."
US-08-655-227-17

Query Match          20.8%; Score 104.2; DB 3; Length 1656;
Best Local Similarity 67.8%; Pred. No. 8.2e-14;
Matches 145; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 10 GGGCTGGGGCTGGGGCTGGAGCTGAGCTGAGCTGAGGCTGGGGCTGGGGCTGG 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1341 GGGCAGCGAGCGAGAGCGAGAGCGAGCTGTGCTGTGCTGTGCTGTGCTGTG 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 GGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1281 CGTCTGAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 AGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1221 AGGTTGTAGCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 GGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1161 GGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-08-655-241-17/c
; Sequence 17, Application US/08655241
; Patent No. 6025543
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.

```

APPLICANT: Weigel, Detlef
TITLE OF INVENTION: Seed Plants Exhibiting Early Reproductive
TITLE OF INVENTION: Development and Methods of Making Same
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,241
FILING DATE: 05-JUN-1996
CLASSIFICATION: CLASS 800
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1894
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1656 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1651
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1656
OTHER INFORMATION: /note= "domain = ecdysone receptor
OTHER INFORMATION: ligand binding domain."
US-08-655-241-17
Query Match 20.8%; Score 104.2; DB 3; Length 1656;
Best Local Similarity 67.8%; Pred. No. 8.2e-14;
Matches 145; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
OY 10 GGGCTGGGCTGGGCTGAGGCTGAGGCTGAGGCTGGGCTGGGCTGGGCTG 69
DB 1341 GGGCAGCGAGCGAGAGCGGAGAGGAGCTGTGCTGTGTAATCTGTGCTGAGACTTG 1282
OY 70 GGGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 129
DB 1281 CGTGTGAGGCTGTGCTGTGAGGCTGGGCTGGGCTGGGCTGGGCTG 1222
OY 130 AGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 189
DB 1221 AGGTTGAGGCTGGGCTGTGTGCTGTGAGATCTTCTGGGTCAGGAGAGGCTTG 1162
OY 190 GGGCTGGGCTGAGGCTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 223
DB 1161 GGGCTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1128
RESULT 5
US-09-398-326-17/c
Sequence 17, Application US/09398326
Patent No. 6355863
GENERAL INFORMATION:
APPLICANT: Yanofsky, Martin F.
TITLE OF INVENTION: Seed Plants Exhibiting Inducible Early
TITLE OF INVENTION: Reproductive Development and Methods of Making Same
NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/398,326
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/659,188
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 3739
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1656 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1651
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..1656
OTHER INFORMATION: /note= "domain = ecdysone receptor
OTHER INFORMATION: ligand binding domain."
US-09-398-326-17
Query Match 20.8%; Score 104.2; DB 4; Length 1656;
Best Local Similarity 67.8%; Pred. No. 8.2e-14;
Matches 145; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
OY 10 GGGCTGGGCTGGGCTGAGGCTGAGGCTGAGGCTGGGCTGGGCTGGGCTG 69
DB 1341 GGGCAGCGAGCGAGAGCGGAGAGGAGCTGTGCTGTGTAATCTGTGCTGAGACTTG 1282
OY 70 GGGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 129
DB 1281 CGTGTGAGGCTGTGCTGTGAGGCTGGGCTGGGCTGGGCTGGGCTG 1222
OY 130 AGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 189
DB 1221 AGGTTGAGGCTGGGCTGTGTGCTGTGAGATCTTCTGGGTCAGGAGAGGCTTG 1162
OY 190 GGGCTGGGCTGAGGCTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 223
DB 1161 GGGCTGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1128
RESULT 6
US-09-144-759-17/c
Sequence 17, Application US/09144759
Patent No. 6117639
GENERAL INFORMATION:
APPLICANT: Hoock, Thomas
APPLICANT: Kwon, Ann
APPLICANT: Hermann, Ursula
TITLE OF INVENTION: FUSION PROTEINS, DNA MOLECULES, VECTORS, AND HOST CELLS
TITLE OF INVENTION: USEFUL FOR MEASURING PROTEASE ACTIVITY

	Query Match	20.8%	Score 104.2	DB 3	Length 2295
	Best Local Similarity	67.8%	Pred. No. 8.7e-14		
	Matches 145	Conservative	0	Mismatches 69	Indels 0
QY	10	GGGCTGGGGCTGGGGCTGAGCTGAGCTGAGCTGGGGCTGGGGCTGGGGCTG	69		
DB	1983	GGGACACGGACGCGAGACAGGAGGAGCGTGGCGTGGTGAATCTTGTGGCTGGAATTG	1924		
QY	70	GGGCTGAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGGGGCTG	129		
DB	1923	CGTCTGAAGCTGTGTGTGTGAGCTGGGGGTTGCAGTTTGACCTTGCAGTGTGAGTGTACTG	1864		
QY	130	AGGCTGGGGCTGGAGCTGAGGCTGGGGCTGGAGCTGGAGCTGGGGCTGGAGGTTG	189		

RESULT 9
US-08-209-747-1
Sequence 1, Application US/08209747
Patent No. 5733771
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNAS Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P. O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Fri Oct 11 09:30:52 2002

us-09-489-101a-17_copy_1_500.rni

Page 8

OY	315	GANGCTG	321
		11	11
Db	304	GATTTTG	310

Search completed: October 10, 2002, 20:05:12
Job time : 24.7333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 695.733 Seconds
(without alignments)
9699.805 Million cell updates/sec

Title: US-09-489-101a-17_COPY_1_500
Perfect score: 500
Sequence: 1 ctgagcgtcggcgtcggcgtc.....gaatnaaatnccaccann 500

Scoring table: IDENTITY_MNC
Gapop 10.0, Capext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	250.8	50.2	616	12	CNS03VH
C 2	250	50.0	723	12	CNS03VH
C 3	247.2	49.4	723	12	CNS02D1N
C 4	237.4	47.5	1061	12	CNS03VH
C 5	237	47.4	880	12	CNS03VH
C 6	237	47.4	949	12	CNS03VH
C 7	226.2	45.2	971	12	CNS04R5
C 8	225	45.0	966	12	CNS02SPN
C 9	224.6	44.9	890	12	CNS03VH
C 10	221.8	44.4	965	12	CNS03VH
C 11	221.6	44.3	899	12	CNS02E22
C 12	219.4	43.9	986	12	CNS03VH
C 13	217.8	43.6	791	12	CNS03VH
C 14	217.8	42.4	1056	12	CNS02R1
C 15	210.2	42.0	1101	12	CNS04L7
C 16	194.6	38.9	906	12	CNS01S6G
C 17	193.4	38.7	475	12	CNS04S6N

C 18	187.8	37.6	731	12	CNS04208
C 19	184.8	37.0	975	12	CNS037S0
C 20	183.8	36.8	959	12	CNS022KX
C 21	183.6	36.7	438	12	AZ858646
C 22	182.6	36.5	805	12	CNS04R42
C 23	167.4	33.5	562	12	AZ855343
C 24	164.6	32.9	812	12	CNS04Q30
C 25	163.4	32.7	485	12	AZ434703
C 26	162.6	32.5	654	12	AZ937136
C 27	162.6	32.5	1091	12	CNS05OEX
C 28	161.6	32.3	614	12	CNS01MDO
C 29	160	32.0	455	12	AQ928308
C 30	160	32.0	832	12	BG481596
C 31	159.8	32.0	1019	12	CNS02HKM
C 32	159	31.8	486	12	CNS01TND
C 33	158.6	31.7	865	12	CNS03ANA
C 34	157.6	31.5	832	12	CNS04FDB
C 35	152.6	30.5	485	9	AA930444
C 36	152	30.4	717	12	BH044236
C 37	150.4	30.1	655	9	BB628062
C 38	149.4	29.9	526	12	CNS03CMO
C 39	149.2	29.8	367	10	BF933221
C 40	148.8	29.8	533	12	AZ958534
C 41	148.6	29.7	961	12	CNS03QKT
C 42	147.6	29.5	837	10	BF697250
C 43	146.6	29.3	335	12	AZ415847
C 44	146.4	29.3	826	12	BH316049
C 45	146	29.2	372	10	BF932849

ALIGNMENTS

RESULT 1
LOCUS CNS03VH/C
DEFINITION Tetraodon nigroviridis genome survey sequence pUC-ori end of clone 062P02 of library G from tetraodon nigroviridis, genomic survey
SEQUENCE
ACCESSION AL262934.1 GI:7984575
VERSION AL262934.1
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 616) Dasilva, C., Fizes, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissbach, J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish tetraodon nigroviridis
JOURNAL Unpublished
AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizes, C., Winkler, P., Brothier, P., Quetier, F., Saurin, W. and Weissbach, J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 616)
AUTHORS Direct Submission
TITLE Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.
FEATURES
Location/Qualifiers
1..616
/organism="Tetraodon nigroviridis"

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Human gene number estimate provided by genome wide analysis using		
2	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Petradon nigroviridis DNA sequence		
3	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Unpublished		
4	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	3 (bases 1 to 613)		
5	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Genoscope.		
6	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Direct Submission		
7	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases		
8	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	This sequence is a single read and was generated as part of a large		
9	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	scale clone-and sequencing project of the Petradon nigroviridis		
10	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	genome. For more information, please take a look at		
11	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	http://www.genoscope.cns.fr/Petradon .		
12	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Location/Qualifiers		
13	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	1. 613		
14	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	/organism="Petradon nigroviridis"		
15	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	/db_xref="taxon:99883"		
16	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	/clone="256821"		
17	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	/clone.lib="G"		
18	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	/note="Genoscope sequence ID : C0AG256C111P1-end : T7"		
19	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	BASE COUNT		
20	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	114 a 123 c 264 g 100 t 12 others		
21	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	ORIGIN		
22	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Query Match		
23	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Best Local Similarity 78.0%; Pred. No. 8.8e-29;		
24	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	Matches 302; Conservative 0; Mismatches 84; Indels 1; Gaps 1;		
25	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	1 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGAGGCTGGGGCTGGGG 60		
26	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	208 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGAGGCTGGGGCTGGGG 267		
27	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	61 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGAGGCTGGGGCTGGGG 120		
28	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	268 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGAGGCTGGGGCTGGGG 327		
29	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	121 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGGGGCTGGGG 180		
30	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	328 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGAGGCTGGGGCTGGGG 387		
31	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	181 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGGGGCTGGGG 240		
32	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	388 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGAGGCTGGGGCTGGGG 447		
33	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	241 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGGGGCTGGGG 300		
34	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	448 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGGGGCTGGGG 507		
35	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	301 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGGGGCTGGGG 359		
36	Journal of Molecular Evolution	1998	Barrois, A., Fzames, C., Mincker, P., Brotter, P., Quetier, F., Saurin, N., and Wesselschlag, J.	508 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGGCTGAGGCTGAGGCTGGGGCTGGGG 567		
37	Journal of Molecular Evolution					

[illegible]

JOURNAL REFERENCE	JOURNAL TITLE	JOURNAL COMMENT	FEATURES	BASE COUNT	ORIGIN
freshwater pufferfish Tetraodon nigroviridis					
Unpublished					
2 (bases 1 to 1056)					
Roest-Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.					
Human gene number estimate provided by genome wide analysts using Tetraodon nigroviridis DNA sequence					
Unpublished					
3 (bases 1 to 1056)					
Genoscope.					
Direct Submission					
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases					
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at					
http://www.genoscope.cns.fr/tetraodon.location/Qualifiers					
1. 1056					
/organism="Tetraodon nigroviridis"					
/db_xref="taxon:99883"					
/clone="167P12"					
/clone_lib="G"					
/note="Genoscope sequence ID : COAG167DC06P1-end : T7"					
190 a 358 c 273 g 234 t 1 others					
Query Match	42.4%	Score 211.8;	DB 12;	Length 1056;	
Best Local Similarity	72.8%	Pred. No. 2.1e-23;			
Matches 278; Conservative	0;	Mismatches 103;	Indels 1;	Gaps 1;	
1 CTGAGGCTGGGGCTGGGGCTGGAGCTGAGCTGGAGCTGGAGCTGGGCTGGG 60					
1006 CCGGGCGAGAGCCGGGGCTAGGCTGAGCTGGGGCTGAGAGCTGAGCTGAGT 947					
61 CTGGGGCTGGGGCTGAGGCTGGGCTGGGGCTGGGGCTGGGGCTGAGCTGAGG 120					
946 CCGAGGCCGGGGCCGAGGCTGAGTCCGAGGCCGGGGCTGAGAGCTGAGCTGAGT 887					
121 CTGGGGCTGAGGCTGGGGCTGGAGCTGAGGCTGGGGCTGGGGCTGGGGCTGGGG 180					
886 CCGAGGCCGGGGCCGAGGCTGAGTCCGAGGCCGGGGCTGAGAGCTGAGGCTGAGT 827					
181 CTGAGTTGGGGCTGGGACTGAGGCTGGGGCTTANGGCTGGGGCTGAGGCTGAGG 240					
826 CTGAGGCTGAGTCCGAGGCCGGGGCCGAGGCTGAGTCCGAGGCTGAGGCTGAGG 767					
241 CTNAGGCTGAGTTGGGGCTGGGCTGGAGCTGAGAGCTGAGGCTGAGGCTGAGG 300					
766 CTGAGGCTGAGTCCGAGGCTGAGGCTGAGTCCGAGGCGGGGCTTGAAGCGGAGCGAGG 707					
301 CTGAGAGCTGGGGCTGAGGCTGGGGCTGGGGCTGAGGCTGAGGCTGAGGCTGAGG 359					
706 CTGAGGCTGAGGCTGAGGCGGGGGCTGAGAGCTGAGTCCGAGGCCGGGGCTGAGGCCGGA 647					
360 CTGAGCTGAGGCTGGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 381					
646 CCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 625					
RESULT 15					
CNS04LE7/c	1101 bp	DNA	linear	GSS 21-MAY-2000	
LOCUS					
DEFINITION	Tetraodon nigroviridis genome survey sequence PUC-Orl end of clone 11B102 of library G from Tetraodon nigroviridis, genomic survey sequence.				
ACCESSION	AL296008.1	GI:8034588			
VERSION					
KEYWORDS	GSS: genome survey sequence.				
SOURCE	Tetraodon nigroviridis.				
ORGANISM	Tetraodon nigroviridis				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;				

REFERENCE	Acanthomorphia: Acanthopterygii; Percomorphia; Tetraodontiformes; Tetraodontidae; Tetraodon.	
AUTHORS	1 (bases 1 to 1101) Roest-Crolius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C., Bounneu,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.	
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 1101) Roest-Crolius,H., Jallion,O., Dasilva,C., Bounneu,L., Fisher,C., Bernot,A., Fitzames,C., Winkler,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.	
AUTHORS	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	
TITLE	Unpublished	
JOURNAL	3 (bases 1 to 1101)	
REFERENCE	Genoscope.	
AUTHORS	Direct Submission	
TITLE	Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases	
JOURNAL	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .	
COMMENT	Location/Qualifiers	
FEATURES	1..1101	
source	/organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="118102" /clone_11b="G" /note="genoscope sequence ID : C0BG118BE01sp1-end : PUC-Or1"	
BASE COUNT	198 a 355 c 282 g 245 t 21 others	
ORIGIN		
Query Match	42.0%: Score 210.2; DB 12; Length 1101;	
Best Local Similarity	74.7%: Pred No. 3; 6e-23;	
Matches	290; Conservative 3; Mismatches 92; Indels 3; Gaps 3	
1	CTGAGCGTGGGCTGGGCTGGGCGCTGAGCGCTGAGCGTGGGACTGAGCGCTGGGG	60
1074	CTGAGGCTGAGGCCGGGGCTGAGGCTGAGTCCGGGGCCGGG-CTGAGGCTRAGGCTRACG	1016
61	CTGGGCGTGGGCTGAGGCTGGGGCTGGGGCTG-GGGCTGGGGCTGGGACTGAGCGTTGG	119
1015	CTGAGCGCGGGGCTGAGTCTGAGTCCGGGGCTGAGGCCAGAGGCTGAGGCCAGAGGCTGAG	956
120	GTGGGGCTGAGCGCTGGGGCTGGGACTGAGAGGCTGGGGCTGGGACTGAGGCTGGGG	179
955	GCGGGGCTGAGGCGCAGGCTGAGGCGCGGGGCTGAGGCTGAGGCTGAGTCCGGGGCTGAG	896
180	GCTGAGTTGGGCGCTGGGACTGAGGCTGGGGCTGANGGCTGGGCTGAGGCTGGGGCTTGG	239
895	GCCGGGCTGAGGCCAGAGGCTGAGGCCGAGGCGTGAAGGCTGAGGCTGAGTCCGGG	836
240	GCTNAGGCTGAGGTTGGGGCTGGGGCTGAGGCTGAGCGTGGGGCTGAGGCTGAGGCTGAG	299
835	GCTGAGGCTGAGGCTGAGGCCGCGGACTGAGTCTGAGTCCGGGGCTGAGGCCGAGGCTGAG	776
300	GCTGAGGCTGGGGCTGANGCTGGGGCTGGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG	358
775	GCCGAGGCTGAGGCCCGGGGCTGAGGCTGAGGCTGAGGCTGAGTCCGGGCTGAGGCTGAG	716
359	GCTGAGGCTGAGGCTGGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG	386
715	GCTGAGGCTGAGGCCGAGGCTGAGTCT	688

Search completed: October 10, 2002, 20:28:37
Job time : 702.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 84.9333 Seconds
(without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101a-17_COPY_1_500

Perfect score: 500
Sequence: 1 ctgagctgctgggctgggct.....gaatnaaatctccacccam 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Genesec 032802: *
2: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1980.DAT: *
3: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1981.DAT: *
4: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1982.DAT: *
5: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1983.DAT: *
6: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1984.DAT: *
7: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1985.DAT: *
8: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1986.DAT: *
9: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1987.DAT: *
10: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1988.DAT: *
11: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1989.DAT: *
12: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1990.DAT: *
13: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1991.DAT: *
14: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1992.DAT: *
15: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1993.DAT: *
16: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1994.DAT: *
17: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1995.DAT: *
18: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1996.DAT: *
19: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1997.DAT: *
20: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1998.DAT: *
21: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA1999.DAT: *
22: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA2000.DAT: *
23: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA2001A.DAT: *
24: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA2001B.DAT: *
25: /SIDSI/gcgdata/hold-genesec/genesec-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	460	92.0	1232	22	AA011124
2	171	34.2	2465	11	AA003665
3	165.8	33.2	564	22	ABA00274
4	165.8	33.2	564	22	ABA28356
5	165.8	33.2	564	22	AAK08552
6	165.8	33.2	564	22	AAK34435
7	165.8	33.2	564	22	AAI16734
8	165.8	33.2	564	22	AAI40161
9	163.4	32.7	569	23	AA584303

10	163.4	32.7	1209	23	AA580389
11	163.4	32.7	1209	23	AA581401
12	161.6	32.3	10347	22	AA728060
13	148.4	29.7	368	22	ABA72813
14	148.4	29.7	368	22	ABA58433
15	148.4	29.7	368	22	AAK21244
16	148.4	29.7	368	22	AAK47401
17	148.4	29.7	368	22	AAI25929
18	148.4	29.7	368	22	AAI53240
19	147.2	29.4	2403	23	ABL24558
20	144.2	28.8	3092	23	AA569141
21	144.2	28.8	10578	23	AA584306
22	143.2	28.6	1548	12	AAQ13868
23	143.2	28.6	303	23	ABL24559
24	142.8	28.6	3526	12	AA013864
25	137.8	27.6	1172	24	AA023835
26	135.8	27.2	716	21	AA044382
27	132.4	26.5	521	21	AA066070
28	132.4	26.5	2701	22	AAH34279
29	131.4	26.2	2548	24	AA023837
30	130.4	26.1	1954	24	AB193802
31	129.2	25.8	486	22	AA77507
32	127.8	25.6	5120	22	AA04677
33	127.4	25.5	6794	21	AA25597
34	125.6	25.1	3780	24	AA52005
35	122.8	24.6	4279	23	AA586600
36	121.4	24.3	2550	19	AAV52497
37	120.8	24.2	6558	21	AA245602
38	119.6	23.9	590	24	AA023836
39	115.4	23.1	2465	11	AA003665
40	114.8	23.0	2668	22	AAK75409
41	114.6	22.9	3975	16	AA092412
42	113.8	22.8	383	16	AA088689
43	113.8	22.8	666	16	AA088690
44	113.8	22.8	807	19	AAV52505
45	113.8	22.8	3633	21	AA05548

ALIGNMENTS

RESULT 1	AA011124	standard; DNA: 1232 BP.
ID	AA011124	
XX	AA011124	
AC	AA011124	
DT	24-SEP-2001	(first entry)
XX		
DE	Human small cell lung cancer associated gene, Novel-2.	
XX		
KN	Human: small cell lung cancer; therapy: hCAAP; nucleic acid: NA;	
KW	melanoma; cancer; colon; breast; head; neck; transitional cancer;	
KM	leiomyosarcoma; synovial sarcoma; cytostatic; ds.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153349-A2.	
PD	26-JUL-2001.	
XX		
PF	19-JAN-2001; 2001WO-US02015.	
XX		
PR	21-JAN-2000; 2000US-0489101.	
XX		
PA	(LUDM-) LUDMG INST CANCER RES.	
PA	(SICK) SIOX KETTERING INST CANCER RES.	
PA	(CORR) CORNELL RES FOUND INC.	
XX		
PI	Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;	
XX		
DR	WPI: 2001-457597/49.	
XX		

DNA encoding novel
Human 07C927 gene
Human foetal liver
Probe #16899 for g
Human brain expres
Human bone marrow
Probe #15862 for g
Probe #15862 used
Drosophila melanog
DNA encoding novel
Insert from clone
Drosophila melanog
ppp 378 gene and r
Sugarane proline
Human secreted exp
Human lung cancer-
Human colon cancer-
Sugarane proline
Mouse ischemic CO
Polyglutamine Eric
DNA sequence of hu
DNA sequence of h
DNA encoding fat-5
DNA encoding novel
Streptococcus pneu
cDNA sequence of a
Sugarane proline
Sequence homology
Human immune/haema
pP1022 4-kb inser
Left terminal pack
Streptococcus pneu
Streptococcus pneu

DE	Human foetal liver single exon nucleic acid probe #8579.
KW	Human: foetal liver: gene expression; single exon nucleic acid probe; ss.
XX	Homo sapiens.
OS	MO200157277-A2.
PN	09-AUG-2001.
PD	30-JAN-2002; 2001WO-US00669.
PE	04-FEB-2000; 2000US-0180312.
XX	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0609408.
PR	03-AUG-2000; 2000US-0632365.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	(MOLE-) MOLECULAR DYNAMICS INC.
PA	Penn SG, Hanzel DK, Chen W, Rank DR;
PI	WPI; 2001-483447/52.
DR	human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human fetal liver -
XX	Claim 1: SEQ ID NO 8579; 639bp + sequence listing; English.
XX	The invention relates to a single exon nucleic acid probe for
CC	measuring human gene expression in a sample derived from human foetal
CC	liver. The single exon nucleic acid probes may be used for predicting,
CC	measuring and displaying gene expression in samples derived from human
CC	fetal liver. The present sequence is a single exon nucleic acid
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part of the
CC	published specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX	Sequence 564 BP; 75 A; 103 C; 227 G; 159 T; 0 other:
SQ	
Query Match	33.2%; Score 165.8; DB 22: Length 564;
Best Local Similarity	64.7%; Pred.No. 9.5e-21;
Matches 264; Conservative	0; Mismatches 142; Indels 2; Gaps 2.
QY	1 CTGAGCGTGGGGGTGGGGGCTGGGGCTGAAGCTTGAGGCTGGAGCTGGAGGCTGGAGG 60
Db	16 CTGTGGCATATGCATGCGACATGTAAGTCGTGGCTGGCGCATTGAGCTGTCACTGACC 75
QY	61 CTGGGGCTGGGGGTGGAGGCTGGGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGG 120
Db	76 CTGTGGCTGTGGCTGTGAATGGGACTGGGGCTGGAGCTATGATGATATGTTCT-TGG 134
QY	121 CTGGGGCTGGAGGCTGGGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGG 180
Db	135 CTGTGTCATATGCTGTAGTCGACATGAGCGCTGGAGGCTGGAGGCTGGAGGCTGGAGG 194
QY	181 CTGAGGTTGGGGGTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGG 240
Db	195 CTATGACGTGTAATGTGTCT-TGGCTGGGCTATGAGGCTGTAGCTGGCACTGTGGCTGTGG 253
QY	241 CTAAAGCTGAGGTTGGGGCTGGGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGG 300
Db	254 CTGTGGCTTCTACTACGCCCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 313
QY	301 CTGAGAGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGGCTGGAGG 360
Db	314 CTCTAATCTGGAGGCTGTAGCTGTGGCTGTGGGCTGTGGGCTGTGGGCTGTGGAGTGGCCCTGGG 373
QY	361 TGAAGCTGAGGCTGGGGCTTAACGCTGAGCTGANNNGCTGGTGGCTNATG 408

Db 374 CTGTGGCTATGACTGCACCGTCCGTGGCTGTGGCTGGCTGGT 421

RESULT 4
ABA28556
ID ABA28556 standard; DNA: 564 BP.
XX
AC ABA28556;
XX
CT 23-JAN-2002 (first entry)
XX
DE Probe #7022 for gene expression analysis in human heart cell sample.
XX
NM Human; gene expression: heart; microarray; vascular system; probe:
KM cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease, ss.
XX Homo sapiens.
XX
XX MO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hazrel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 1; SEQ ID No 7022; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging, and
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 564 BP; 75 A; 103 C; 227 G; 159 T; 0 other;

Query Match 33.2%; Score 165.8; DB 22; Length 564;
Best Local Similarity 64.7%; Pred. No. 9, 5e-21;
Matches 264; Conservative 0; Mismatches 142; Indels 2; Gaps 2

Db 16 CTGTGGCTATGACTGCACCGTCCGTGGCTGTGGCTGGCTGGT 75

Y 61 CTGGGCGCTGGGCGCTGAGCTGGGCGCTGGGCGCTGGGCGCTGGGCGTGAAGCTGGG 120

Y 76 CTGTGGCTATGACTGCACCGTCCGTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGG 134

Y 121 CTGGGCGCTGGGCGCTGAGCTGGGCGTGAAGCTGGGCGCTGGGCGTGGGCGTGGGCG 180

Fri Oct 11 09:30:51 2002

us-09-489-101a-17_copy_1_500.rng

Page 10

OY		61	CTGGAGGTGAGGCCTGAAGCCTTGGAGCCATGGCGTCGACCTTGAGGCCAAGACTGTTGGG	120
Dd		76	CTGGTGTATTGGCTGTAGCTTAGCTGACCATTGTGGCTGTGTAGATGGCACTGTGGCTGTGG	135
OY		121	CTGGGGCTAGAAGCTTGGGGCTTGGGACAATAAGCTTGGGCTTGGCACTAGAAGCTTGGGCTGGG	180
Dd		136	CTATTGACTGTATATGTTTTGG-CTGATCTTAAGCTTAACTTAACTGGCACATGTGGCTGTGG	194
OY		181	CTGAGCTTGGGGCTGGGACTAGAAGCTTGGGCTTANAGGCTTGGGGCTTGAAGGCTTGGGCTAAGG	240
Dd		195	CTTGGGCTCTGACTTGGCCCCCTGTGGCTGTGGCTGTCTCTCTGTGGATGTAACTTGGCACTGTGG	254
OY		241	CTTNAGCTGATAGGTTGGGGCTTGMNGCTGACGCTTGGGCTTGAAGCTTGGGCTTGNAGCAGG	300
Dd		255	CTTGAAGCTGTGCTGTGACCTTGAGCTGTGGCTGTGCTGTGTGTGCTGTGACTTGGCCCTGTGTG	314
OY		301	CTTGAGCTTGGGGCTTGANMGCTTGGGGCTTGGGCTTNNMNCCTTAMNCTTGGGCGCTGA	351
Dd		315	CTGTGGCTTATGACTTGCCACCTTGGCCCCTGTGGCTGTGGCTGTGTGTGTGTGA	365

RESULT 15
ARK21244 standard; DNA; 368 BP.

XX AC	ANK21244;
DT DT	05-NOV-2001 (first entry)
DE DE	Human brain expressed single exon probe SEQ ID NO: 21235.
OS OS	Human; brain expressed exon; gene expression analysis; probe:
KW KW	microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW KW	epilepsy; cancer; ss.
XX XX	Homo sapiens.
PB PB	NM0200157275-A2.
PD PD	09-AUG-2001.
XX XA	30-JAN-2001; 2001WO-US00667.
FF FF	04-FEB-2000; 2000US-O180312.
PR PR	26-MAY-2000; 2000US-O207456.
PR PR	30-JUN-2000; 2000US-O608408.
PR PR	03-AUG-2000; 2000US-O632366.
PR PR	21-SEP-2000; 2000US-O234687.
PR PR	27-SEP-2000; 2000US-O236359.
PR PR	04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human brains -

Example 4: SEQ ID NO: 21235; 650bp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.

Sequence 368 BP; 29 A; 76 C; 149 G; 114 T; 0 other:

Query Match	29.7%	Score 148.4	DB 22	Length 368
Best Local Similarity	66.7%	Pred. No. 1e-17		
Matches 224	Conservative 0	Mismatches 115	Indels 2	Gaps 2
QY	1	CTGAGGCTGCGGCGCTGCGGCGCTGAGGCGTGAAGCTGGAGGACTGAGGCGTGGGCGTGGG	60	
Db	17	CTGTGGCTGTGGCTGTGTAGTGCACGTGGGCTGTGGGCTATGAGGCTGATGCTGTGT-7GG	75	
QY	61	CTGGGGCTGGGGGCTGAGGCTGGGGGCTGGGGGCTGGGGGCTGGAGCTGAGGCTGGGG	120	
Db	76	CTGGCTCATGTGGCTGTAGCTGTGCACCTGTGGCTGTGGCTGTGAGTGACACTGTGGCTGGG	135	
QY	121	CTGGGGGCTGAGGCGTGGGGGCTGGGACCTGAGAGGCTGGGGGCTGGAGGCTGTGGGGCGTGGGG	180	
Db	136	CTATGACTGTGTACTGTGTGTCTTTGG-CTGGTGCTATAGGCTGTGAGCTGTGGCACTGTGGCTGTGG	194	
QY	181	CTGAGGTTGGGGCTGGGACCTGAGGCGTGGGGCTTAMGGCTGGGGGCTGAGGCTGGGGGCTGAGG	240	
Db	195	CTGTGGCTCTGTGACTGGGCGCCGTGTGGGCTGTGGGCTGTGGGCTGTGGGATGTAGCACTGGCACTGGG	254	
QY	241	CTTAAAGCTTGAAGTTGGGGGCTGGGGGCTGGGAGCTGAAGCTGGGGGCTGAGAGCTTGMAGCTGAGG	300	
Db	255	CTGTACTGTGGCTGTGTGACTGTGTGTGTGGCTGTGGCGTGGAGGCTGTGACTGTGGCCCTGGGG	314	
QY	301	CTTGAAGCTGGGGGCTGAGAGCTGGGGGCTGGGGGCTGGGAGGCTGAGAGCTGAGAGCTGGGGGCTGA	351	
Db	315	CTTGGGCTATGTACTGTGCACCGTGGGCGTGGGCTGTGGGCTGTGGGCTGTGGTGTGGTGA	365	

Search completed: October 10, 2002, 15:18:11
Job time : 91 secs

Fri Oct 11 09:30:50 2002

us-09-489-101a-17_copy_1_500.rge

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 715.867 Seconds

(without alignments)
14616.225 Million cell updates/sec

Title: US-09-489-101a-17_COPY_1_500

Perfect score: 500
Sequence: 1 ctgagcgtcggcgtcggcgtc.....gaatnaaatcaccacann 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Capext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genbank: 1: gb_da: 2: gb_hg: 3: gb_in: 4: gb_lm: 5: gb_ov: 6: gb_pat: 7: gb_pl: 8: gb_pl: 9: gb_pl: 10: gb_pl: 11: gb_pl: 12: gb_pl: 13: gb_pl: 14: gb_pl: 15: gb_pl: 16: gb_pl: 17: gb_pl: 18: gb_pl: 19: gb_pl: 20: gb_pl: 21: gb_pl: 22: gb_pl: 23: gb_pl: 24: gb_pl: 25: gb_pl: 26: gb_pl: 27: gb_pl: 28: gb_pl: 29: gb_pl: 30: gb_pl: 31: gb_pl: 32: gb_pl: 33: gb_pl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	460	92.0	1232	6	AX201596	Sequence
2	289.4	57.9	54987	2	AC094895	Rattus no
3	285	57.0	143720	10	AC087332	Mus Muscu
4	280.4	56.1	220761	2	AL670227	Mus muscu
5	258.4	51.7	2801	5	AF218784	Gallus ga
6	248.8	49.8	172224	2	AC087149	Mus muscu
7	225	45.0	118562	9	AC026443	Homio sapi
8	225	45.0	134909	2	AC025183	Homio sapi
9	220.8	44.2	179386	3	AC104515	Drosophi1
10	219.4	43.9	182421	3	AC104515	Drosophi1
11	217.6	43.5	179595	2	AC098110	Rattus no
12	215	43.0	2773	10	AF189263	Mus muscu
13	215	43.0	110312	10	AF131866	Mus muscu
14	214.8	43.0	7085	2	AC015260	AE003574
15	214.8	43.0	274388	3	AC020884_3	Drosophi1
16	213.8	42.8	73038	2	AC020884_3	Drosophi1
17	213.4	42.7	1935	10	AF189264S2	Continuatio
18	209.8	42.0	122592	3	CEY60A3A	Continuatio
19	208.8	41.8	39498	2	CEY60A3A	Continuatio
20	205.4	41.1	3000	10	AF189262	Continuatio
21	205.4	41.1	290452	9	AC079167	Continuatio
22	204.2	40.8	139934	2	HUB384B8	Continuatio
23	203.4	40.7	3978	10	AF255385	Continuatio
24	199.8	39.6	165849	2	AC096270	Continuatio
25	197.8	39.0	192202	2	AC094105	Continuatio
26	195.2	39.0	167224	2	AC095838	Continuatio
27	194.4	38.9	67391	2	AC101718	Continuatio
28	193.6	38.7	148871	2	AC025324	Continuatio
29	193.6	38.6	3768	9	AC015261	Continuatio
30	191.8	38.4	146339	2	HUM8AC	Continuatio
31	191.4	38.3	154825	2	AC103031	Continuatio
32	189.6	37.7	34727	3	CEP11A6	Continuatio
33	188.6	37.4	88370	2	AC084153	Continuatio
34	186.2	37.2	226502	2	AL645802	Continuatio
35	186.2	37.2	226502	2	AL645802	Continuatio
36	185.4	37.1	257817	2	MOSH5Y1B	Continuatio
37	185.4	36.9	66491	2	AC006909	Continuatio
38	184.4	36.7	191692	2	AC006715	Continuatio
39	183.6	36.7	157909	2	TERKSTELA	Continuatio
40	183.6	36.6	157909	2	AC023834	Continuatio
41	183	36.3	34727	2	AC096043	Continuatio
42	181.4	36.2	2679	3	CEY11A6	Continuatio
43	179.6	35.9	43255	2	CEY11A6	Continuatio
44	179.6	35.9	43255	2	CEY11A6	Continuatio
45	179.6	35.9	43255	2	CEY11A6	Continuatio

ALIGNMENTS

RESULT 1	AX201596	1232 bp	DNA	linear	PAT 30-AUG-2001
LOCUS	AX201596				
DEFINITION	Sequence 17 from Patent WO015349.				
ACCESSION	AX201596				
VERSION	AX201596.1	GI:15391446			
KEYWORDS					
SOURCE					
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Shockett, E., Scanlan, M.J., Jager, D., Old, L.J., Gure, A.O. and				
	Chen, Y.F.				
TITLE	Small cell lung cancer associated antigens and uses therefor				
JOURNAL	Patent: WO 015349-A 17 26-Jul-2001.				
	LUDWIG. INSTITUTE FOR CANCER RESEARCH (US) ; MEMORIAL				
	SLOAN-KETTERING CANCER CENTER (US) ; CORNELL RESEARCH FOUNDATION,				
	INC. (US)				
FEATURES					
Source	1..1232	Location/Qualifiers			
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				

Db	48624	C7GAGGCTGAGGCTGAGGCTGAGGCTGAGGCGCTAAGGAGCTGAAGCCTGGGCGACTGAG	48565
Oy	241	C7MACGCTGAGGTTGGGCGTGGGCGTGAGMOCCTGACGCCTGAGGCTGAGGCTMGAGMCTGAG	300
Db	48564	C7GGGGCTGAGGTCTGGGGCTGAGGCGCTGGGCGCTGAGGCTGAGGCTGAGGCTGGGCGCTGAG	48505
Oy	301	C7GAGGCTGGGGGCTGAGGCTGGGGGCTGGGGGCTGGNNCTGA-NOTGGGGCTGAGGCTCCN	359
Db	48504	C7GAGGCTGGGGGCTGAGGCTGGGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGGGG	48445
Oy	360	C7GAGGAGGAGGCTGGGGGCT-----TMAAGCTGAGCTGNNNGCTGCTCTAAATCTTG	412
Db	48444	C7GAGGCTGAGGCTGGGGGCTGGGGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGG	48387
RESULT 3			
LOCUS	AC087332	143720 bp	DNA linear ROD 25-JAN-2002
DEFINITION	Mus musculus chromosome 2 Rp33-58L18, complete sequence.		
ACCESSION	AC087332		
VERSION	AC087332.6	GI:12642989	
KEYWORDS	HTG;		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 143720)		
REFERENCE	Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R. High Throughput Mouse Sequencing Unpublished 2 (bases 1 to 143720)		
TITLE	Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R. Direct Submission Submitted (28-DEC-2000) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave., Bronx, NY 10461, USA 3 (bases 1 to 143720)		
REFERENCE	Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R. Direct Submission Submitted (01-FEB-2001) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave., Bronx, NY 10461, USA 4 (bases 1 to 143720)		
REFERENCE	Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R. Direct Submission Submitted (04-JAN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St., Cambridge, MA 02139, USA 5 (bases 1 to 143720)		
REFERENCE	Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R. Direct Submission Submitted (25-JAN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St., Cambridge, MA 02139, USA On Feb 1, 2001 this sequence version replaced gi.12484303.		
JOURNAL	Center: Harvard Partners Genome Center Web site: http://wchanning.bwh.harvard.edu:9086/hpcpg/jsp/hpcpg/Sequence/mouse.html Contact: gnikm@csapecod.bwh.harvard.edu		

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550). Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as low coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards. Estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as low quality.

Summary Statistics

Center project name: AJP
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 1008
Assembly program: Phrap version 0.990319
Contig length: 143720
Fraction of Phrap value < 40: 8.349e-05
Error rate in Consed: 0.00 per 10,000 bases
Number of N's in consensus: 0

Distribution of Quality < 40 Bases:

Quality	Count
10001	1
9001	1
8001	1
7001	1
6001	1
5001	1
4001	1
3001	1
2001	1
1001	1
01	1

Phrap Value Range	Count
5	1
10	1
15	1
20	1
25	1
30	1
35	1
40	1

FEATURES

source
Location/Qualifiers
1. 143720
/organism="Mus musculus"
/db_xref="taxon:10090"
repeat_region
2373..2464
/rpt_family="L3"
complement(3181..3264)
repeat_region
3265..3380
/rpt_family="L3"
complement(3381..3819)
repeat_region
3819..3914
/rpt_family="L3"
complement(3915..4014)
repeat_region
4015..4014
/rpt_family="B4A"

Query Match	Score	DB	Length
repeat_region	complement(4015..4232)		
repeat_region	/rpt_family="B3A"		
repeat_region	complement(4233..4449)		
repeat_region	/rpt_family="B4A"		
repeat_region	4529..4567		
repeat_region	/rpt_family="Cn-rich"		
repeat_region	complement(4636..4779)		
repeat_region	/rpt_family="B1_MM"		
repeat_region	complement(4957..5065)		
repeat_region	/rpt_family="L1M4"		
repeat_region	7137..7206		
repeat_region	/rpt_family="PB1D10"		
repeat_region	7208..7231		
repeat_region	/rpt_family="(CAAA)n"		
repeat_region	7462..7906		
repeat_region	/rpt_family="(GGA)n"		
repeat_region	8548..8632		
repeat_region	/rpt_family="B1_MM"		
repeat_region	10774..10795		
repeat_region	/rpt_family="(CAGAGA)n"		
repeat_region	complement(10946..11091)		
repeat_region	/rpt_family="B1_MM"		
repeat_region	11502..11580		
repeat_region	/rpt_family="(TG)n"		
repeat_region	12288..12443		
repeat_region	/rpt_family="RSINE1"		
repeat_region	complement(12521..12648)		
repeat_region	/rpt_family="MIR"		
repeat_region	complement(15305..15523)		
repeat_region	/rpt_family="MIR"		
repeat_region	17262..17351		
repeat_region	/standard_name="456819"		
repeat_region	complement(19924..20048)		
repeat_region	/rpt_family="B1F"		
repeat_region	26227..26603		
repeat_region	/rpt_family="(CTGGGG)n"		
repeat_region	27851..27887		
repeat_region	/rpt_family="(CA)n"		
repeat_region	27887..27921		
repeat_region	/rpt_family="(GA)n"		
repeat_region	29704..29733		
repeat_region	/rpt_family="(TA)n"		
repeat_region	30264..30401		
repeat_region	/rpt_family="B1_MM"		
repeat_region	33524..33686		
repeat_region	/rpt_family="B2_MM1"		
repeat_region	33687..33709		
repeat_region	/rpt_family="(TAAA)n"		
repeat_region	33878..34603		
repeat_region	/standard_name="D2Arz18"		
repeat_region	complement(35888..36002)		
repeat_region	/rpt_family="PB1D7"		
repeat_region	36558..36822		
repeat_region	/rpt_family="L1M4"		
repeat_region	36830..36938		
repeat_region	/rpt_family="PB1D7"		
repeat_region	36962..37038		
repeat_region	/rpt_family="L1M4"		
repeat_region	38396..38754		
repeat_region	/rpt_family="MTD"		
repeat_region	39284..39405		
repeat_region	/rpt_family="PB1D10"		
repeat_region	39406..39431		
repeat_region	/rpt_family="(GGA)n"		
repeat_region	39448..39650		
repeat_region	/rpt_family="B3A"		
repeat_region	39771..40266		
repeat_region	/rpt_family="L1M4"		
repeat_region	40273..40321		
repeat_region	/rpt_family="(TG)n"		
repeat_region	complement(40705..40894)		

57.0%; Score 285; DB 10; Length 143720;

[illegible]

Db	357	G6GCTGGGGCGTGAGGCATGGGGGTCTGGGCTGGGGCTGGGGCTG		416
Oy	316	ANGCTTGGGCTGGGGCGCANNCCGTG	339	
Db	417	G8GCTGGGCTGGGGCTGCACGTG	440	

RESULT 6			
LOCUS	AC087149		
DEFINITION	Mus musculus clone Rp23-31010, WORKING DRAFT SEQUENCE, 17 unordered pieces.		
ACCESSION	AC087149	172224 bp	DNA linear HTG 09-DEC-2000
VERSION	AC087149.1	GI:11610875	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eumaxillops; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 172224)		
JOURNAL	DOE Joint Genome Institute.		
REFERENCE	Sequencing of Mouse		
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 172224)		
REFERENCE	DOE Joint Genome Institute.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (09-DEC-2000) Production Sequencing Facility, DOE Joint		
COMMENT	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
	-----Genome Center		
	Center: Joint Genome Institute		
	Center Code: JGI		
	Web site: http://www.jgi.doe.gov		

	Project Information		
	Center Project Name: 1856110		
	Center clone name: RPCI_23_31016		

Summary Statistics			
Consensus quality: 160964 bases at least Q40			
Consensus quality: 164326 bases at least Q30			
Consensus quality: 166482 bases at least Q20			
Estimated insert size: 200000; agarose-gel estimation			
Estimated insert size: 170624; sum-of-contigs estimation			
Quality coverage: 10.97 in Q20 bases; agarose-gel estimation			
Quality coverage: 12.86 in Q20 bases; sum-of-contigs estimation			
NOTE: This is a working draft sequence. It currently			
consists of 17 contigs. The true order of the pieces			
is not known and their order in this sequence record is			
arbitrary. Gaps between the contigs are represented as			
runs of N, but the exact sizes of the gaps are unknown.			
This record will be updated with the finished sequence			
as soon as it is available and the accession number will			
be preserved.			
*	1	1216:	contig of 1216 bp in length
*		1316:	gap of unknown length
*	1217	1316:	gap of unknown length
*	1317	2574:	contig of 1258 bp in length
*	2575	2674:	gap of unknown length
*	2675	3683:	contig of 1009 bp in length
*	3684	3783:	gap of unknown length
*	3784	6495:	contig of 2712 bp in length
*	6496	6595:	gap of unknown length
*	6596	9535:	contig of 2940 bp in length
*	9536	9635:	contig of unknown length
*	9636	11503:	contig of 1868 bp in length
*	11504	11503:	gap of unknown length
*	11604	15360:	contig of 3757 bp in length
*	15361	15460:	gap of unknown length
*	15461	26040:	gap of 10580 bp in length
*	26041	26180:	gap of unknown length
*	26181	34081:	contig of 7941 bp in length
*	34082	34181:	gap of unknown length
*	34182	41397:	contig of 7116 bp in length
*	41398	41397:	gap of unknown length

JOURNAL

Submitted (13-DEC-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

COMMENT

This sequence was assembled using end and sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

FEATURES

location=0qualifiers

source 1. 179386
/organism="Drosophila melanogaster"
/strain="y: cn bw 99"
/db_xref="taxon:7227"
/map="20B-20C"
/clone="BACR27K09 (D1451)"
/clone_lib="RPCL198 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial Bcont1 in
PBAS2.6)"
BASE COUNT 53055 a 35376 c 36862 g 54093 t
ORIGIN

Query Match 44.2% Score 220.8; DB 3; Length 179386;
Best Local Similarity 67.7% Pred. No. Be-29; Indels 1; Gaps 1;
Matches 302; Conservative 0; Mismatches 143;
1 CTGAGCGTGGGCGCTGGGCGTGAAGCTGGAGCTGAGCTGGGCGTGGG 60
2252 CTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 2311
61 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 120
2312 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2371
121 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 180
2372 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2431
181 CTGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 240
2432 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2491
241 CTGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 300
2492 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2551
301 CTGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 359
2552 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2611
360 CTGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 419
2612 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2671
420 ATGNNATGNNCTGNNCTGNNCTGNNCTGNNCTGNNCTGNNCTGNNCTGNN 445
2672 CTGGAGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 2697

RESULT 10
AC011760 182421 bp DNA linear HTG 18-FEB-2000
LOCUS AC011760
DEFINITION Drosophila melanogaster chromosome X clone BACR0522 (D1132)
RPL1-98 05.K.22 map 20C-20C strain y: cn bw sp. *** SEQUENCING IN
PROGRESS ***
AC011760
AC011760.8 GI:6996986
KEYWORDS HTG; PHASE1.

SOURCE

fruit fly.
Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 182421)
Celniker, S.E., Agbayani, A., Arcana, T.T., Baxter, E., Blazek, R.G.,
Butenkov, C., Chape, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,
Moshrefi, A.R., Nixon, K., Paclob, J.M., Park, S., Pfeiffer, B.,
Richards, S., Sethi, H., Svirskas, R.R., Man, K.H., Webster, D.,
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.

AUTHORS

TITLE

2 (bases 1 to 182421)
Celniker, S.E., Agbayani, A., Arcana, T.T., Baxter, E., Blazek, R.G.,
Butenkov, C., Chape, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,
Moshrefi, A.R., Nixon, K., Paclob, J.M., Park, S., Pfeiffer, B.,
Richards, S., Sethi, H., Svirskas, R.R., Man, K.H., Webster, D.,
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.

AUTHORS

TITLE

Submitted (13-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Feb 18, 2000 this sequence version replaced gi:6980168.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the 100% identity, 200 base length, 200 base overlap criteria.

COMMENT

NOTE: This is a working draft sequence. It currently
consists of 116 contigs. The true order of the pieces
is not known and the order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N. Below the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved. 583: contig of 583 bp in length
1 584: gap of unknown length
* 664: gap of unknown length
* 1132: contig of 449 bp in length
* 1133: gap of unknown length
* 1135: contig of 624 bp in length
* 1817: gap of unknown length
* 1897: contig of 437 bp in length
* 2334: gap of unknown length
* 2414: gap of 523 bp in length
* 2536: contig of 523 bp in length
* 2537: gap of unknown length
* 3017: contig of 403 bp in length
* 3420: gap of unknown length
* 3499: gap of 565 bp in length
* 4064: contig of 565 bp in length
* 4144: gap of unknown length
* 4065: gap of unknown length
* 5070: contig of 926 bp in length
* 5071: gap of unknown length
* 5151: contig of 1253 bp in length
* 6404: gap of unknown length
* 6484: contig of 645 bp in length
* 7129: gap of unknown length
* 7209: gap of 660 bp in length
* 7209: contig of unknown length
* 7869: gap of unknown length
* 7948: gap of 516 bp in length
* 8464: contig of unknown length
* 8544: gap of unknown length
* 9096: contig of 552 bp in length
* 9176: gap of unknown length
* 9177: contig of 551 bp in length
* 9728: gap of unknown length
* 9808: contig of unknown length
* 10889: contig of 1082 bp in length
* 10890: gap of unknown length

```

10970 11838: contig of 869 bp in length
11839 11918: gap of unknown length
11919 12546: contig of 628 bp in length
12547 12626: gap of unknown length
12627 14048: contig of 1422 bp in length
14049 14128: gap of unknown length
14129 14941: contig of 813 bp in length
14942 15021: gap of unknown length
15022 15912: contig of 891 bp in length
15913 15992: gap of unknown length
15993 16730: contig of 738 bp in length
16731 16810: gap of unknown length
16811 18632: contig of 1822 bp in length
18633 18712: gap of unknown length
18713 20042: contig of 1330 bp in length
20043 20122: gap of unknown length
20123 21959: contig of 1837 bp in length
21960 22039: gap of unknown length
22040 22782: contig of 743 bp in length
22783 22862: gap of unknown length
22863 23949: contig of 1087 bp in length
23950 24029: gap of unknown length
24030 25382: contig of 1353 bp in length
25383 25462: gap of unknown length
25463 27227: contig of 1765 bp in length
27228 27307: gap of unknown length
27308 28704: contig of 1397 bp in length
28705 28784: gap of unknown length
28785 30730: contig of 1946 bp in length
30731 30810: gap of unknown length
30811 32148: contig of 1338 bp in length
32149 32228: gap of unknown length
32229 34377: contig of 2149 bp in length
34378 34457: gap of unknown length
34458 36419: contig of 1962 bp in length
36420 36499: gap of unknown length
36500 37371: contig of 872 bp in length
37372 37451: gap of unknown length
37452 39687: contig of 2536 bp in length
39688 40067: gap of unknown length
40068 42766: contig of 2639 bp in length
42767 43846: gap of unknown length
43847 44960: contig of 2350 bp in length
44961 45276: gap of unknown length
45277 47714: contig of 2458 bp in length
47715 47834: gap of unknown length
47835 50642: contig of 2788 bp in length
50643 50722: gap of unknown length
50723 54355: contig of 3633 bp in length
54356 57368: contig of 2933 bp in length
57369 57448: gap of unknown length
57449 61101: contig of 3653 bp in length
61102 61181: gap of unknown length
61182 66754: contig of 5573 bp in length
66755 66834: gap of unknown length
66835 68984: contig of 2150 bp in length
68985 69064: gap of unknown length
69065 73317: contig of 4233 bp in length
73318 73397: gap of unknown length
73398 76661: contig of 3264 bp in length
76662 76741: gap of unknown length
76742 80988: contig of 4248 bp in length
80989 81069: gap of unknown length
81070 86081: contig of 5012 bp in length
86082 86161: gap of unknown length
86162 90536: contig of 4375 bp in length
90537 90616: gap of unknown length
90617 97771: contig of 7155 bp in length
97772 97851: gap of unknown length
97852 105505: contig of 7654 bp in length
105506 105585: gap of unknown length
105586 113220: contig of 7635 bp in length

```

```

113221 113300: gap of unknown length
113301 122668: contig of 9368 bp in length
122669 122748: gap of unknown length
122749 143365: contig of 20617 bp in length
143366 143445: gap of unknown length
143446 144051: contig of 606 bp in length
144052 144131: gap of unknown length
144132 144542: contig of 411 bp in length
144543 144622: gap of unknown length
144623 145227: contig of 605 bp in length
145228 145307: gap of unknown length
145308 145957: contig of 650 bp in length
145958 146037: gap of unknown length
146038 146243: contig of 206 bp in length
146244 146902: gap of unknown length
146903 146982: contig of 579 bp in length
146983 147298: gap of unknown length
147299 147378: contig of 316 bp in length
147379 148089: gap of unknown length
148090 148169: contig of 711 bp in length
148170 148859: gap of unknown length
148860 148939: gap of unknown length
148940 149370: gap of unknown length
149371 149450: contig of 431 bp in length
149451 150063: contig of 613 bp in length
150064 150143: gap of unknown length
150144 150284: contig of 441 bp in length
150285 150664: gap of unknown length
150665 151269: contig of 605 bp in length
151270 151350: gap of unknown length
151351 151837: contig of 588 bp in length
151838 152017: gap of unknown length
152018 152643: contig of 628 bp in length
152644 152723: gap of unknown length
152724 153396: contig of 671 bp in length
153397 153476: gap of unknown length
153477 154126: contig of 650 bp in length
154127 154206: gap of unknown length
154207 154679: contig of 473 bp in length
154680 154759: gap of unknown length
154760 155063: contig of 304 bp in length
155064 155143: gap of unknown length
155144 155454: contig of 311 bp in length
155455 155534: gap of unknown length
155535 156215: contig of 680 bp in length
156216 156294: gap of unknown length

```

Query Match
 Best Local Similarity 71.3%; Score 219.4; DB 2; Length 182421;
 Matches 291; Conservative 0; Mismatches 116; Indels 1; Gaps 1;

```

QY 1 CTGAGGCTGGGGCTGGGGCTGGGGCTGGAGCTGGAGCTGGAGCTGGGGCTGGGG 60
DB 110544 CTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGGA 110603
QY 61 CTGGGGCTGGGGCTGGAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGAGCTGGAG 120
DB 110604 CTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGGA 110663
QY 121 CTGGAGCTGGAGCTGGGGCTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGGGCTGGGG 180
DB 110664 CTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGGA 110723
QY 181 CTGAGCTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 240
DB 110724 CTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGGA 110783
QY 241 CTNAGCTGGAGCTGGGGCTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 300
DB 110784 CTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGGA 110843
QY 301 CTGAGCTGGGGCTGGAGCTGGGGCTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 359

```


repeat_region	/rpt_family="(TAAA)n"	complement(23395). 23446)
repeat_region	/rpt_family="(GA)n"	23448. 23483
repeat_region	/rpt_family="(CA)n"	complement(23547). 23581)
repeat_region	/rpt_family="(TAAA)n"	complement(23582). 23739)
repeat_region	/rpt_family="A2"	complement(23809. 23949)
repeat_region	/rpt_family="TTR1B"	complement(23950. 25074)
repeat_region	/rpt_family="T11_MM"	complement(25075. 25527)
repeat_region	/rpt_family="TTR1A"	complement(26021. 26112)
repeat_region	/rpt_family="MTR45B"	27955. 28172
repeat_region	/rpt_family="B4A"	28227. 28247
repeat_region	/rpt_family="AT_rich"	complement(28855. 28891)
repeat_region	/rpt_family="(CA)n"	complement(29584. 29770)
repeat_region	/rpt_family="B3"	complement(31305. 31467)
repeat_region	/rpt_family="RMR1A5"	complement(31573. 31624)
repeat_region	/rpt_family="(GA)n"	31626. 31710
repeat_region	/rpt_family="(CA)n"	complement(33260. 33276)
repeat_region	/rpt_family="L1"	332765. 33198
repeat_region	/rpt_family="L1_MM"	complement(33509. 33739)
repeat_region	/rpt_family="L2"	complement(35041. 35122)
repeat_region	/rpt_family="B1_F"	complement(35274. 35320)
repeat_region	/rpt_family="purine-rich"	35347. 35424
repeat_region	/rpt_family="polypurine"	complement(35431. 35540)
repeat_region	/rpt_family="RSTN1"	complement(35613. 35797)
repeat_region	/rpt_family="LX8"	complement(35863. 36017)
repeat_region	/rpt_family="LX"	complement(36150. 36257)
repeat_region	/rpt_family="PB1D10"	complement(36363. 36702)
repeat_region	/rpt_family="B4"	complement(36709. 36769)
repeat_region	/rpt_family="A-rich"	complement(36770. 36921)
repeat_region	/rpt_family="B1_MM"	37088. 37110
repeat_region	/rpt_family="(CA)n"	complement(37142. 37347)
repeat_region	/rpt_family="B4"	complement(37367. 37470)
repeat_region	/rpt_family="L2"	37570. 37656
repeat_region	/rpt_family="MT1B"	37746. 37836
repeat_region	/rpt_family="MTR1B"	complement(37865. 38302)
repeat_region	/rpt_family="L1_MM"	38270. 38667
repeat_region	/rpt_family="L1_MM"	38671. 38684
repeat_region	/rpt_family="MT1B"	

Fri Oct 11 09:30:50 2002

us-09-489-101a-17_copy_1_500.rge

Page 17

Dh
182 CGACGACGACGACGACGACGACGACGACGACGACGACGACGA 241

[illegible]

```

RESULT 2
US-09-103-840A-2/c
: Sequence 2, Application US/09103840A
: Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO. 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

[illegible]

Db 3937747 ATGCGCGCGCGCGCGCTTGCCCGCGGTCGTCGTCGCGCCAGC 3937703
 RESULT 3
 US-09-197-649-7/c
 ; Sequence 7, Application US/09197649
 Patent No. 6194550

```

1 TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Transla
2
3 FILE REFERENCE: NEX02/CI-CON
4
5 CURRENT APPLICATION NUMBER: US/09/197,649
6
7 CURRENT FILING DATE: 1998-11-23
8
9 EARLIER APPLICATION NUMBER: 07/829,461
10
11 EARLIER FILING DATE: 1992-01-31
12
13 EARLIER APPLICATION NUMBER: 07/779,055
14
15 EARLIER FILING DATE: 1991-08-01
16
17 EARLIER APPLICATION NUMBER: 07/551,968
18
19 EARLIER FILING DATE: 1990-08-02
20
21 NUMBER OF SEQ ID NOS: 26
22
23 SOFTWARE: PatentIn Ver. 2.0
24
25 SEQ ID NO 7
26
27 LENGTH: 390
28
29 TYPE: LNA
30
31 ORGANISM: Artificial Sequence
32
33 FEATURE:
34
35 OTHER INFORMATION: Description of Artificial Sequence: Sequence
36
37 OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
38
39 OTHER INFORMATION: fragments having NcoI restriction sites.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
99
```

[illegible]

Fri Oct 11 09:30:47 2002

us-09-489-101a-16_copy_1_500.rni

Page 3

```

1  APPLICANT: Sutcliffe, Thomas D.
2  APPLICANT: Rodriguez, Raymond L.
3  TITLE OF INVENTION: Production of Mature Proteins
4  TITLE OF INVENTION: In Plants
5  NUMBER OF SEQUENCES: 23
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: Dehlinger & Associates
8  STREET: 350 Cambridge Ave., Suite 250
9  CITY: Palo Alto
10 STATE: CA
11 COUNTRY: USA
12 ZIP: 94306
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Diskette
15 COMPUTER: IBM Compatible
16 OPERATING SYSTEM: DOS
17 SOFTWARE: FastSeq for Windows Version 2.0
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/09/023.173
20 FILING DATE: 13-FEB-1998
21 CLASSIFICATION: 435
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: 60/038.168
24 FILING DATE: 13-FEB-1997
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Pelithory, Joanne R
27 REGISTRATION NUMBER: P4295
28 REFERENCE/DOCKET NUMBER: 0665-0007.30
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 650-324-0880
31 TELEFAX: 650-324-0960
32 INFORMATION FOR SEQ. ID NO.: 4:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 1140 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38 IMMEDIATE SOURCE:
39 CLONE: Codon-optimized Ramy3D signal-prosubtilisin BPN'
40 US-09-023-173-4

```

[illegible]

Oy	425	CGCACCCAGGAGATGTGCACAGGGCATCCTCACAGGACGCCGCCA	484
Db	690	CCCGAGCGCAGCGCGCCCTCAAGGCGCGCTGCACAAGGCGTCGCAGCGGCTGT	749
Oy	485	GCGCATCGCC	494
Db	750	CGTCTCGCC	759

RESULT 5
US-09-056-556-176
Sequence 176, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yael A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-056-556-176

	Query Match	Similarity	15.4%	Score 77	DB 4	Length 494
	Best Local	Similarity	49.6%	Pred. No. 3.8e-06		
	Matches 127	Conservative	0	Mismatches 200	Indels	Gaps
QY	38	CGCATATGAGGACGCGCCGACAGGGGATTCGCTATGAGAGACACACACCATGCTGATGGCCAA	97			
DB	16	CGGCAGAGGGCGGCGACAGCGTGGCCACAGCGGCGTACGGCGGTGGGGGCAACAATGTTCCGTAC	75			
QY	98	CGAGGAAAGCGCCGACGGGACATGCGCGAGAGAGCCATCCAGAGGATGCGCAACGAGAGGT	157			
DB	76	CGAAGCGGCGACCGCGCGGTGGCGCGCGCCGCGCGCGCCGCGGCGAGCGGCGTTTTCGG	135			
QY	158	TGCCCGAGGGCATGCGCAATGGGTGGTGGCCACAGGCGCATGCCAATGAGAGACGCCACCA	217			
DB	136	CGGAGAGAGGGCGGCTTCGAGCGGCGACGGCGCGGTAGAGCGCGGCCCAACAGCGGCGGTACCGT	135			
QY	218	GGGCGATGCGCAACTGGGACGCGCCCTCCACGCGTTCTGGCCACAGGGGAGACCGCGTCTTACGTT	277			
DB	196	CGGATCCCGTGGCCGGTGGCGGCGGACGACGCGCGGTCTCGGCGCGCGGGGGGAGAGGGGCT	255			
QY	278	CGCCAAACGGGAGCGCGCCGACGAGGACATCGCCACACGGGAGGCCAACCGAGGCGATGGGCA	337			
DB	256	CTTTGCGCGGTGCGCGGCGAGCCAGGCGCGAGCTCGGTGGTGGAGGCGGCGCAATGGTGGCGGTCTC	315			

Fri Oct 11 09:30:47 2002

us-09-489-101a-16_copy_1_500.rni

Page 4

OY	338	CAGAGTACCATTCACAGGATACGCTTACGAGAGACGCGTCCAGGATCGTACAGAGT	397
Db	316	CACCGCGCGGACAGGCGGCGCTTTGGGCGGCGCGGCGCGTGTGCGAGGACACGCCCGCGCTCG	375
OY	398	GGCGCGCCAGGCGATCGTGGCCACGAGACGCGCCGAC	434
Db	376	TGCGGAATCCGGGCTACCATCGAGACACGCGGCGCAG	412

```

RESULT 6
US-09-130-114-2/c
: Sequence 2, Application US/09130114
: Patent No. 5976807
: GENERAL INFORMATION:
: APPLICANT: Horlick, Robert A.
: APPLICANT: Demaj, Bassam B.
: APPLICANT: Robbins, Alan K.
: TITLE OF INVENTION: Eukaryotic Cells Stab
: TITLE OF INVENTION: From Multiple Transfe
: FILE REFERENCE: 0867/ID903051
: CURRENT APPLICATION NUMBER: US/09/130,114
: CURRENT FILING DATE: 1998-08-06
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSPD for Windows Version 3.0
: SEQ ID NO: 2
: LENGTH: 1931
: TYPE: DNA
: ORGANSIM: EBNA
US-09-130-114-2

```

Query Match	15.4%;	Score 76.8;	DB 2;	Length 1931;
Best Local Similarity	47.7%;	Pred. No. 4.2e-06;		
Matches 225;	Conservative 0;	Mismatches 247;	Indels 0;	Gaps 0

[illegible]

RESULT 7
US-09-056-556-167
; Sequence 167, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 1439 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Query Match	Similarity	15.2%	Score	76.2	DB.4:	Length	1439:
Best Local	Similarity	48.9%	Pred.	No. 5.3e-06:			
Matches	225:	Conservative	0:	Mismatches	243:	Indels	3:
							Gaps
							1
QY	22	ACGCTGCCAAGGAGCATGCTAATGAGACAGCCGACACGAGGATCGCTAATGAGAGACCA	81				
Db	937	ACCCGAGCAGCAGCACACCGGACAGCGCGCTCCGCGTGGCGGGGTATATGAGGGGCACTG	996				
QY	82	CCCAATGATGATGGCCCAAGAGAAAGCC---GCCCAAGGATGCGCCGAGAGGCAATCCAG	138				
Db	997	GCTCTAAGCGGCGGGGGGGGTCTGTGGCGGGGCGCGGGAACGGGGGTGTCTGCGCGGT	1056				
QY	139	GCATTCGCCAAGAGAGATTGGCCAGGGATATCGCCATATGGGATTCGCCGACAGGGCATCG	198				
Db	1057	CCCTTCGCAACGCTGTGTGGGCGGGAGACGGCGGCACAGCGGTGTAACGGGGGCAAGGGCGCG	1116				
QY	199	CGAATATGAGACGCCACCCAGGAGCATGGCCATATGGAGAGCGCTGTCCAGGCTTTCGCAAG	258				
Db	1117	ACGGCAGCAGCGGCGCGCGCCGCGGCAAGGGCGGCACAGGCAAGCGGTATCGCCGCAAGC	1176				
QY	259	GGGAGCGCGTCTCTACGCTTTCGCCAAGGGGGAAGCGCCCAAGGCGCATTCGCCAAGGGGAGC	318				
Db	1177	GCTCAAGGCTGTCTCAAGCTCACACCGCCGCGCCACAGGGAGGGAACGGGGGCAATATGGCGCAAC	1236				
QY	319	CCACCAAGGCGATGGGCAACGAGAGTACCATCCATCGCATTCCTAACGAGAGACGCCGTCC	378				
Db	1237	GCGGCGACAGCGCTCCGCGCGGCGCGGAGGCGCAAGGGGATGGCGGCGGACCGCGGCCAAGC	1296				
QY	379	AGGGCATTCGCTAATGAGATGAGCCGCGCCAGGAGCATTCGCCAACGAGAGCGCCGCCAGGGAA	438				
Db	1297	GCGGCGACAGCGCGGGGTGGCCACAGCGGGGCGCCAGAGGGGTAAAGGGCGGCCAAGCGCACAGC	1356				
QY	439	TCCGCGAGGATATGTGCACAGAGGATCGCCCAACGAGAGACCGCCGACAGGGCATTCGCCAACA	498				
Db	1357	GTTCCGCGCAGAGGGGTCAAGAGGCTCATCAACGTACCGCGCGGCGCAAGCGGCTGCAACAGCGCGCA	1416				
QY	499	A	499				

Db 1417 A 1417

RESULT 8

US-09-105-537-7
: Sequence 7, Application us/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, H.
: TITLE OF INVENTION: DNA encoding methymycin and plitromycin
: FILE REFERENCE: 600,438051
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 1248
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-105-537-7

Query Match 14.9% Score 74.4; DB 4; Length 1248;
Best Local Similarity 49.5%; Pred. No. 1,le-05;
Matches 220; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 59 GGGGATCCCTAATGAGACACCAACGATGATCCGCAACGAGAGCCGCCAGGGCAT 118
Db 444 GGGGATCCCTAATGAGACACCAACGATGATCCGCAACGAGAGCCGCCAGGGCAT 503
QY 119 CGCCGAGAGAGCCATCCAGGACATGCGCAACGAGAGGTTGCCAGGGCATGCGCAATGG 178
Db 504 GTACTTCGACGCGCGGACGCGCTCGGCTGCGGCGGTGCGAGCGCGCGCGCGCGCGCT 563
QY 179 GGTGCGCCACAGGAGGATCCGCAATGAGAGAGGAGGAGGAGGATCCGCAATGGAGAGC 238
Db 564 CGCGGAGAGGAGGATCCGCAATGAGAGGAGGAGGAGGAGGATCCGCAATGGAGAGC 623
QY 239 CGTTCACGCGCTTCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 295
Db 624 CGCGGAGAGGAGGATCCGCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
QY 296 CGAGGAGATCCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 355
Db 684 CTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 743
QY 356 CATTCGATACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 415
Db 744 CGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 803
QY 416 CAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 475
Db 804 CAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 863
QY 476 CGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 499
Db 864 CGACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 887

RESULT 9
US-09-056-167/C
: Sequence 167, Application us/09056556
: Patent No. 6350456
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Skelky, Yasir A.W.
: APPLICANT: Dillon, David C.
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
: NUMBER OF SEQUENCES: 241
: CORRESPONDENCE ADDRESS:

TREATM

ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-Apr-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 1439 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-056-167

Query Match 14.9% Score 74.4; DB 4; Length 1439;
Best Local Similarity 47.6%; Pred. No. 1,le-05;
Matches 219; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 8 CGGATTCCTAGAGACCTCTCCAGGCGATCCCTAATGAGAGAGCCGACAGGCGATCCG 67
Db 1413 CGCGGATTCCTAGAGACCTCTCCAGGCGATCCCTAATGAGAGAGCCGACAGGCGCG 1354
QY 68 TAATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 127
Db 1353 TGTGTGCGCTTCGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1294
QY 128 CGCATTCAGGAGGATCCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 187
Db 1293 TGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1234
QY 188 ACAGGAGATTCGCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 247
Db 1233 TGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1174
QY 248 CTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 307
Db 1173 TGTGCGACCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1114
QY 308 CAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 367
Db 1113 CGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1054
QY 368 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 427
Db 1053 CGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 994
QY 428 CGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 467
Db 993 TGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954

RESULT 10
US-09-105-537-3/C
: Sequence 3, Application us/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zuo, J.
TITLE OF INVENTION: DNA encoding methymycin and plikromycin
FILE REFERENCE: 600 438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 3
LENGTH: 13613
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-3

Query Match 14.9%; Score 74.4; DB 4; Length 13613;
Best Local Similarity 49.5%; Pred. No. 1.3e-05;
Matches 220; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 59 GGGCATGCTAATAGAGACACCCAGTGCATGCCAAGAGAGAAAGCCGAGGCGCAT 118
DB 12440 GGGCCGCTCTGGCCGCGGACAGCTGGAGAGTGGCCAGAGAGCGGCTGGGCT 12381
QY 119 CGCCGAGAGAGCCATCCAGGCGATGCCAAGAGAGAGTGGCCAGGCGCATGCGCATG 178
DB 12380 GTACTTGGAGCGCGCGGACGCTCGGCTGGCGGTGGAGCGCGCGCGGAGGCT 12321
QY 179 GGTGGCGGACAGGAGGATGCCAATAGAGAGAGCGGCGATGCCAAGTGGAGAGCG 238
DB 12320 CGGCGAGCGCGGAGGCTTCAAGTTCAGCGGACGAGGCGGCTGACGCGCTGAGGCGG 12261
QY 239 CGTCCAGCGGCTTGGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 295
DB 12260 CGCGGTGTACGAG 12201
QY 295 CCAGGCGATGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 355
DB 12200 CTTCGACTGCTCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 12141
QY 356 CATGCTAATAGAGAGAGCGGCTCCAGGAGATGCTAAGAGTGGCGCGGCGGAGGAGGAG 415
DB 12140 CGCGGCGATGGGCTTCACTCTCCGAGCGGCTTCCGAGGTATCGAGCGGAGAGCGGAG 12081
QY 416 CAACGAGAGCGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 475
DB 12080 CAACGAGAGCGCGGCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 12021
QY 476 CGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 499
DB 12020 CGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 11997

RESULT 11
US-09-177-349-2/c
Sequence 2, Application US/09177349
Patent No. 6268201
GENERAL INFORMATION:
APPLICANT: Alland, David
APPLICANT: Bloom, Barry R.
APPLICANT: Jacobs, Jr., William R.
TITLE OF INVENTION: 11B INHA AND INIC GENES OF MYCOBACTERIA AND METHODS
FILE REFERENCE: 96/00/491
CURRENT APPLICATION NUMBER: US/09/177,349
CURRENT FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentia Ver. 2.0
SEQ ID NO 2
LENGTH: 5036
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
US-09-177-349-2

Query Match 14.8%; Score 74; DB 4; Length 5036;
Best Local Similarity 48.2%; Pred. No. 1.3e-05;
Matches 209; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 7 ACGGATGCTTGGAGAGACGCTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 66
DB 935 ACTGCGTGTGCTGAGACCGCGGTGAGGCGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 876
QY 67 CTAAATGAGACACCTACCCAGTGCATGCCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 126
DB 875 CGGCTGAGACACCTACCCAGTGCATGCCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 816
QY 127 ACGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 186
DB 815 CCAGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 756
QY 187 CACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 246
DB 755 CGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 696
QY 247 GCTTCGCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 306
DB 695 CCAGGAG 636
QY 307 CCAGGAG 366
DB 635 CCATATCTTACCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 576
QY 367 AGGAG 426
DB 575 CCAGGAG 516
QY 427 CGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 440
DB 515 AGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502

RESULT 12
US-09-320-878-21/c
Sequence 21, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: MCDANIEL, Robert
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentia Ver. 2.0
SEQ ID NO 21
LENGTH: 5970
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-320-878-21

Query Match 14.8%; Score 74; DB 3; Length 5970;
Best Local Similarity 49.3%; Pred. No. 1.3e-05;

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 12588 base pairs

```

FILE OF INVENTION: DNA
 FILE REFERENCE: 600.43805

TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051

```

1 CURRENT APPLICATION NUMBER: US/09/105,537A
2 CURRENT FILING DATE: 1998-06-26
3 NUMBER OF SEQ ID NOS: 43
4 SOFTWARE: FASTSEQ for Windows Version 3.0
5 SEQ ID NO 1
6 LENGTH: 15872
7 TYPE: DNA
8 ORGANISM: Streptomyces venezuelae
9 US-09-105-537-1

```

Query Match	14.6%	Score 73.2;	DB 4;	Length 15872;
Best Local Similarity	47.2%;	Pred. No. 1.8e-05;		
Matches 222; Conservative	0;	Mismatches 248;	Indels 0;	Gaps 0;

QY	8	CGCAGATTCGTGAGGACGGCTGCCCGAGGGACATCGCTAATGAGAACCCGACACAGGGACATCGC	67
Db	4792	CGCCATGCGGCGAGGAGAACTGCAGCGCGACCCCGTGTTCGGCGCGCGCTCGACACGCT	4851
QY	68	TAAATGAGGACACACCACCGATGATCGACACAGAGAACGCCGCCAGGGCATTCGCCAGAGA	127
Db	4852	GTTACGGGGCCCTTCGACCGCTCACCTTGACCGCGCGCTGGGGAGATCGTGGCGCGCGGAGA	4911
QY	128	CGCCTATCCAGGGCATTCGCCAACGAGAGTTCGCCAGGGCATTCGCCAATGAGGTTCGCCCC	187
Db	4912	GAGAGCTGAGACCTCACCGGCTGTACACCCAGCCCGCCCTTCCTCCATCGAGGTGGCGCTGTT	4971
QY	188	ACAGGGCATTCGCGCATTCAGGACGCCACCCAGGGCATTCGCCAATCGGGAGCCGCTCCACGG	247
Db	4972	CCGCTCTCTCGAAGCACGACGCCCTCGTCCCGGACGTGTCAACGGGCCATCTGTCGGGCA	5031
QY	248	CTTCGGCCAAAGGGGAGCGCCGCTTCCTCABCTTCGCCAAAGGGGAGGCCGCCAGGAGCATTCGC	307
Db	5032	GATCGCCGCGCGCGCACAGTGCCTGGTGTCTCTTCCTCGACAGAGCGCGACATGTCCTGCAC	5091
QY	308	CAACGGGAGCGCCACCAAGGGCATGGGACATGGATCCACATCCACGGCATTCGCTAAACA	367
Db	5092	CGCCCGCGGCGCGGCTCAATGACAGTGGCGCCGCGAGAGGCGCGCGATGATTCGCCCTGCAGGC	5151
QY	368	GGACGCGCTTCAGAGGATTCGCTAAACGAGTGGCGGCCGCCAGGGCATTCGCCAAACGAGACGC	427
Db	5152	GGGCGAGGCGCGAGGTGCTCGAGTGTCCGTAAGGGGTACAGAGGGCATGGGTTGCGCGTCCGCC	5211
QY	428	CGGCCGAGGATTCGCGAGGATTCGTCACAGGGCATTCGCCAAGAGGAGAGC	477
Db	5212	CGTCAACGAGACCTCACCGCTGTGTCTTCGCGACAGCGAGGACCGCGCGC	5261

RESULT 15
US-09-025-691-2
Sequence 2, Application US/09025691
Patent No. 6069299
GENERAL INFORMATION:
APPLICANT: Broadway, Roxanne M.
APPLICANT: Herman, Gary R.
TITLE OF INVENTION: FOODS AND INSECT CONTROL WITH
TITLE OF INVENTION: CHITINOLYTIC ENZYMES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,691
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,737
REFERENCE/DOCKET NUMBER: 19603/20120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1294 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA

Query Match	14.5%;	Score 72.6;	DB 3;	length 1294;
Best Local Similarity	52.0%;	Pred. No. 2.3e-05;		
Matches 193; Conservative	0;	Mismatches 169;	Indels 9;	Gaps 1

QY	128	CGCCATCCAGGAGCATGCCAAGAAGAGATTTGGCCAGAGGCATATCCCAATATGGGTGGCCG	187
Db	317	CACCGGCCCCGACACCGCGGAGAGGGGGCCCCCGCGGCCAGGCCTCCCGGCGCGGT	376
QY	188	ACAGGGCATCGCCAAATGAGAGAGCGACCACAGGGCATATCGCAATCGAGGAGCGAGCGTCACAGG	247
Db	377	ACCGGCGCCAGCGGTGACCGGTTACTGTGCGAGAACTTCAACAAGGGCGCAGCGTGGDAAAC	436
QY	248	CTTTCGCCAAGGGGGAGCGCGTCTCACCCTTTCGCCAAGGGGAGCGCGCCCGAGGGCATGC	307
Db	437	CCTCGCGAGGTGCGGGAGCGCTACAGCATCATCATGCGCGTCTCTTTCGCGAGCGCCACGCG	496
QY	308	CAAGGGGAGCGCCACCAAGGGGATGGGACAGAGATGCACATTCAGCGGATGGTGTATACGA	367
Db	497	CACCGGGGGAATACCTTACCTTCAGCTGGGTGGGGCTCGCGGTACACCGAGAGA	556
QY	368	GGACCGCGTCCAGGGGATCGCTAACAGAGTGGGCGCGCCAGGGCATTCGGCCACAGAGACGC	427
Db	557	G-----CAATTCCGCGCGACCTTCGCGCCGCAACGACGACGACGCAATGTGGTAT	607
QY	428	CGCCGAGGAAATGCGCGAGATGTTCGACAGGGCATTCGCCAACAGAGATCGCGCCACAGG	487
Db	608	CATCTCGTGGCGGGGAGAGAGGGGCGGTGCGCTCAACACAGACGCTTCGCGCCACGG	667
QY	488	CATGCGCCAA	498
Db	668	CTTCGCGACA	678

Search completed: October 10, 2002, 20:05:06
Job time : 1144.73 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 695.733 seconds
(without alignments)
9699,805 Million cell updates/sec

Title: US-09-489-101a-16_copy_1_500

Sequence: 1 gagctacgcatgctgtag.....cccaagggcatgcacacaaag 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: EST.*

1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estmnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gD_estl:*
10: gD_estl2:*
11: gD_hic:*
12: gD_gsa:*
13: em_gsa_hum:*
14: em_gsa_hnv:*
15: em_gsa_pin:*
16: em_gsa_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153.4	30.7	697	9	AL110383 DKF2P434E
2	124.6	24.9	179	9	AL041090
3	118.6	23.7	758	10	BI829297 603079427
4	113	22.6	673	10	BI561135 603253662
5	98.6	19.7	582	10	BG724320 602697968
6	94.2	18.8	241	9	AI904151 CM-BT043-
7	85.4	17.1	780	10	BG847089 102401580
8	80.8	16.2	822	12	CNS04SGK
9	79.6	15.9	640	10	BE427000 PSR6561 I
10	79.2	15.7	884	10	BG417341 HVSMEK001
11	76.5	15.3	920	9	BE216934 EST0477 T
12	76.4	15.0	681	12	AV176645 AV176645
13	74.8	14.8	915	12	CNS03P82
14	73.4	14.7	511	10	AV438522 AV438522
15	73.4	14.0	511	10	BG739444 EMI_81_A0
16	73.4	14.0	511	10	BG739444 EMI_81_A0
17	70.2	14.0	511	10	BG739444 EMI_81_A0

18	69.8	14.0	454	9	BE125156
19	69.6	13.9	673	10	BG907710
20	69.2	13.8	538	10	BM178969
21	69.2	13.8	679	9	AL506812
22	69.2	13.8	814	10	BE260740
23	68.6	13.7	554	10	BG367845
24	68	13.6	530	10	BP483260
25	68	13.6	689	9	AU076299
26	67.8	13.6	619	10	BG300721
27	67.6	13.5	564	10	BP619259
28	67.6	13.5	652	10	BG314481
29	67.4	13.5	321	10	BG874723
30	67.4	13.5	938	9	BE213813
31	67.2	13.4	625	10	BG906607
32	67.2	13.4	925	12	CNS0092P
33	67.2	13.4	704	9	AU088578
34	66.8	13.4	461	10	BE317259
35	66.8	13.3	456	9	AU088571
36	66.6	13.3	529	10	BE474993
37	66.6	13.3	585	10	BE470936
38	66.6	13.3	676	9	AV916390
39	66.6	13.3	633	9	AV941861
40	66.4	13.2	596	10	BE360540
41	66.2	13.2	770	10	BE265823
42	66.2	13.2	459	10	BE605180
43	66	13.1	401	9	AV438522
44	65.6	13.1	537	10	BG904964
45	65.6	13.1	537	10	BG904964

ALIGNMENTS

RESULT 1
LOCUS AL110383
DEFINITION DKF2P434E0731.t1 434 (synonym: htes3) Homo sapiens cDNA clone
AL110383
ACCESSION AL110383.1 GI:5865991
VERSION DKF2P434E0731.5, mRNA sequence.
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Blum, et al.)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Blum H
MIPS
Am Klopferstr. 18a D-82153 Martinsried, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No si sequence available.
This clone (DKF2P434E0731) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.
Location/Qualifiers
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2P434E0731"
/clone_id="434 (synonym: htes3)"
/tissue="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSPORT1; Site_1: NotI; Site_2: SalI"

FEATURES

BASE COUNT 217 a 181 c 180 g 119 t

Dn 68 TCGTTTAAACAGCATTCGCAATAGCTGGTGCTGTATTATGCATGCCAATAATGACGAATCTGT 127
||||||| | ||| | | ||||| | | | | | ||||| |

Oy 242 CCACGGCTTCCCAAAGGGGAGCACCCTCPACGTTCCAGCAAGGGAGCACGCCGCCAGG 301
||||||| | | | | | | | | | | | | | | | | | | |

Dd 128 ATATGACTATCACTTAATGAGAGCGCTCTATPAGCATCATCATTAATGAGAACCGCTGTATCAAA 187
||||||| | | | | | | | | | | | | | | | | | | |

Oy 302 CATGCCACAAGGGGAGGAGCCACCAAGGGATGGGGAACAGSAGTCACCATCCAGSCATATCGC 361
||||||| | | | | | | | | | | | | | | | | | | |

Dd 188 CATTCCCTAATATGAGACCGCTGTATTATGSGATCGCTCAATGASAGATGCTGTATACAAATTCCG 247
||||||| | | | | | | | | | | | | | | | | | | |

Oy 362 TAGACAGAAGCCGCTCCAGGCGATTCGCTAAGAGTAGTGGCCCCCAGGCATTCGCCAACGA 421
||||||| | | | | | | | | | | | | | | | | | | |

Dd 248 TAATAGATGCTCTTATTATGACTTCCTAATGAGAGCACATCTGTACAGACATCTGTAAAAA 307
||||||| | | | | | | | | | | | | | | | | | | |

Oy 422 GGAGCCCGC 430
|| | |

Dd 308 AGAAGATGC 316

RESULT 4
B1561135 LOCUS B1561135 673 bp mRNA linear EST 05-SEP-2001
DEFINITION 60325366cF1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296262 5'',
MARG sequence.
ACCESSION B1561135
VERSION B1561135.1 GI:15448449
KEYWORDS EST.

ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 673)
NIH-MGC http://mhc.ncbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MG) Unpublished (1999)

AUTHORS CONTACT: Robert Strausberg, Ph.D.
JOURNAL Email: crabs-femail.nih.gov
COMMENT Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CONA Library Preparation: Michael J. Brownstein (NHGRI), Shiraiki Toriyaki and Piero Garinchi (RKEN)
CONLA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution by: MGC Clone Distribution Information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>

Plasmid Map Plate: LHAM1749 row: k column: 15
High quality sequence stop: 630.
Location/Qualifiers
I. 673
source
FEATURES
location
organism="Homo sapiens"
db_xref="taxon:9606"
clone_image="5296262"
clone_lib="NIH_MGC_97"
lab_host="DH10B"
note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgagca); Oligo-dt primed using primer 5'-tttttttttttytvn-3' and size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Garinchi, In preparation). Library constructed by M. Brownstein (NHGRI/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT
172 a 185 c 131 t

Query Match Score 113: DB 10: Length 673:
Best Local Similarity 71.2% Pred No. 1.1e-06,
Matches 163: conservative 0; mismatches 65; Indels 1; Gaps 1,

199 CCATGAGGAGCCACCAAGGCGATCGCACTGGAGCGCCGTCACAGGCTTCCGACAG 258
||||||| | | | | | | | | | | | | | | | | | | |

DB	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	443	CCGATGGGAGACCATCATCAGCATACGACAAAGGAGAGCGCGTCCAGCAGCAGCAGCAGC	502					
0y	259	GGGAGACCGCTTCCATAGCTTCCGCAAGGGGAGAGCGCGCCAGAGGAGCATCGCAGAGGAGC	318					
Db	503	AGGAGCGCGCTCATATGACACCGCCGAGAGAGAGCGCGTCCAGGAGCATCGCGCAGAGAGC	562					
0y	319	CCGACCAAGGAGGATGGGAGAGGATGACCATTCACAGGAGCATCGCTAACGAGGAGCGCTTC	378					
Db	563	CCCTCCAGGAGGATCGTCACTGGAGAGCGCC-TCCATGTCTCTCGCTGACGAGAGAGCGCCCA	621					
0y	379	AGGAGCATCGCTAACGAGATGCGCGCCGAGAGCATTCGCCAAGAGAGCGC	427					
Db	622	CGGCTTCGCTGTATGAGGAGAGCCAGCCAGGAGCATCGCTAACAGGAGAGC	670					
RESULT 5								
LOCUS	B6724320	582 bp	mRNA	linear	EST-08-MAY-2001			
DEFINITION	602697968f1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:482993 5',							
ACCESSION	B6724320							
VERSION	B6724320.1	GI:14003507						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
TITLE	NIH-MGC http://mgi.nci.nih.gov/.							
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)							
COMMENT	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: cgaaps-remail.nih.gov							
	Tissue Procurement: Miklos Palkevics, M.D., Ph.D.							
	CDNA Library Preparation: Michael J. Brownstein (NMGRI), Shiraki							
	Toshiyuki and Piero Carninci (RIKEN)							
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)							
	DNA Sequencing by: Invitae Genomics, Inc							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LLNL at:							
	http://image.llnl.gov							
	Place: LBLM10/50 row: g column: 18							
	High quality sequence stop: 360.							
FEATURES								
source	Location/Qualifiers							
	1..362							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:482993"							
	/clone_lib="NIH_MGC_97"							
	/lab_host="DH10B"							
	/note="Organ: testis; Vector: pBluescript (modified							
	pBluescript KS+); Site_1: BamHI; Site_2: SalI; XhoI (gtcgag							
); Oligo-dT primed using primer 5'-tttttttttttttttttt-3',							
	size-selected for average insert size 2.2 kb and							
	normalized to ROT 5. This is a primary library enriched							
	for full-length clones and constructed using the							
	Cap-trapper method (Carninci, in preparation). Library							
	constructed by M. Brownstein (NMGRI/NHRI, National							
	Institutes of Health). Note: this is a NIH_MGC Library."							
BASE COUNT	183 a 147 c 151 g 101 t							
ORIGIN								
Query Match	19 7%: Score 98.6: DB 10: Length 582:							
Best Local Similarity	66.6%: Pzed No. 9e-05:							
Matches 203: Conservative	0: Mismatches 94: Indels 8: Gaps 4:							
0y	29	CCAGGCGCATCGCTATAGAGAGCCGACGAGGAGCATCGCTAACGAGAGAGCAGT	88					
Db	18	CCGAGGATCGCGACAGAGAGCGCGCCCA-GGCTATGCCAAGAGAGCGCGCCCA-G	75					
0y	89							

QY	329	CATGG	333
QY	149	CGAGGAGGTTGCCAGGAGATGCGCATATGGGGTGC CGCACAGGCGCATGCGCCATATGAGA	208
Db	135	CGGGGACACCTCAACGCGCTGCTTAACAAATACACTGACACACACATGCTTAATGAGG	194
QY	209	CGCACACCAGGACATGCGCCAACTGGGAGCCGCTGACAGGGTTCGCCAACGGGAGAGCGCGT	268
Db	195	CACCTACAAAGATCATACCAAGTAGGGGGCTTTATAGAGAACTTGCTAATAGGAGCCGAC--	252
QY	269	CTCAGACTTTCGCAAGGGGAGCGCGCCGAGGAGATGCCAACGGGGAGCGACACCAAGG	328
Db	253	---AAGGACGCTACTGAGACCTGTACAGACACTTGTCTAATGAGGACAGCGTTATAGA	309
QY	329	CATGG	333
Db	310	CATGG	314

RESULT	6
A1904151	
LOCUS	241 bp mRNA linear EST 30-MAR-2000
DEFINITION	CM-BT043-090299-075 BT043 Homo sapiens cDNA, partial sequence.
ACCESSION	A1904151
VERSION	A1904151.1 GI:6494538
KEYWORDS	EST,
SOURCE	human.

REFERENCE
AUTHORS

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G.

FEATURES
SOUR

BASE COUNT	76 a	42 c	59 g	64 t
ORIGIN				
Query Match	18.8%	Score 94.2;	DB 9;	Length 241;
Best Local Similarity	62.6%	Pred. No. 0.00038;		
Matches 147; Conservative	0;	Mismatches 86;	Indels 0;	Gaps 0;

QY	192	GGCATCGCATTAAGGACGCGCACCCAGGACGATGCGCAACATGGGAGCCCGCTGCACGGCTTC	251
QY	193		
Db	7	GACATCGCATTAAGGACGCGTTGATATGGACATTCCTTAATGAGGATGCTTTACAAAGACATA	66
QY	252	GGCAACGGGGAGCGCGCTCCACAGTTTCCGCAAGGGAGACGCGCCGACGGGCGTCCCAAC	311
QY	253		
Db	67	GGCATATGAGTGTGCTGATATATGATCATCGCTATATGACATTTATATATGACATCTATAT	126
QY	312	GGGGAGCGCCACCAAGGACATGGGCAACGATGCATCTCCAGCGATCGCTTAAGAGAGAC	371
QY	313		
Db	127	GAGAGCGCTCATATGCGATCGCTATATGAGACGCTGTATACACATCTCCCTTAATGAGGAC	186
QY	372	GGCGTCCAGGCGCATCGCTAACGAGGTGGCGCCGACGGGACATGCCCAACGAGAGAG	426
QY	373		
Db	187	GGTGATATGGCATCGCTAATGAGATGCTGTATACGAATTCGCTTATATTAAGGATG	241

RESULT 7			
LOCUS	BG847089		
DEFINITION	BG847089	780 bp	mrna linear
ACCESSION	1024015H06.y2	C. reinhardtii CC-1690, normalized,	EST 29-MAY-2000
VERSION	BG847089	Chlamydomonas reinhardtii cDNA, mRNA sequence.	lambda zap II
KEYWORDS	BG847089.1	GI:14228373	
	EST.		

REFERENCE

JOURNAL COMMENT
unpublished (2000)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000

```

Query Match      17.1%; Score 85.4; DB 10; Length 780;
Best Local Similarity 50.3%; Pred. No. 0.0051;
Matches 240; Conservative 0; Mismatches 231; Indels 6; Gaps 1.

0Y 16 CTGAGAGACCTGCCAGGGCATCCTATTAGAGAGCCCGACGAGCATCCTATTAGAG 75
||||| ||||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 220 CTGACGACCTGTGACGACGCTCCGCTGTGAGAAAGGCGGCCCTCCGCAAGGCGCTCCGCA 279

Qy 76 AC 135

Db 280 AGAAG 339

Qy 136 AGGAGCATCCGCAACGAG 195

Db 340 ACAGAGCCCCGAG 399

Qy 196 TCGCCATGAG 255

Db 400 GCGAC 459

Qy 256 ACAGGAG 315

Db 460 AGGAG 519

Qy 316 ACAGGAG 375

Db 520 AGT 576

Qy 376 TCGAG 435

Db 577 ---AGGAG 633

Qy 436 GAATGCGGAG 492

Db 634 ACAGACACGATGTGACGAG 690

RESULT 8
CNS04SG/c 822 bp DNA linear GSS 24-MAY-2000
LOCUS Tetracodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 007M14 of library H from Tetracodon nigroviridis, genomic survey
sequence.
ACCESSION AL305165.1 GI:8196096
VERSION AL305165.1
KEYWORDS GSS: genome survey sequence.
SOURCE Tetracodon nigroviridis.
ORGANISM Tetracodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetracodon.
REFERENCE 1 (bases 1 to 822)
AUTHORS Roest-Crollius H., Jallion O., Dasilva C., Fizames C., Fisher C.,
Bonneau L., Billault A., Quetier F., Sautin W., Bernot A. and
Weissenbach J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetracodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 822)
AUTHORS Roest-Crollius H., Jallion O., Dasilva C., Bonneau L., Fisher C.,
Bernot A., Fizames C., Wincker P., Brothier P., Quetier F.,
Sautin W. and Weissenbach J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetracodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 822)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetracodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetracodon.
FEATURES
SOURCE
1..822
/organism="Tetracodon nigroviridis"
/db_xref="taxon:99883"
/clone="007M14"

BASE COUNT 147 a 262 c 203 g 209 t 1 others

ORIGIN

Query Match 16.28; Score 80.8; DB 12; Length 822;
Best Local Similarity 50.98; Pred. No. 0.021; Indels 3; Gaps 1;
Matches 217; Conservative 0; Mismatches 206

Qy 20 GAGAGCTGCGGAG 79

Db 556 GAG 497

Qy 80 CACAG 136

Db 496 AGGAG 437

Qy 137 GAG 196

Db 436 AGGAG 377

Qy 197 GAG 256

Db 376 GAG 317

Qy 257 GAG 316

Db 316 GAG 257

Qy 317 GAG 376

Db 256 GAG 197

Qy 377 GAG 436

Db 196 GAG 137

Qy 437 GAG 492

Db 136 GAG 690

RESULT 9
BE427000 640 bp mRNA linear EST 24-JUL-2000
LOCUS PSR6561 ITPC PSR wheat Endospem library Triticum aestivum cDNA
DEFINITION Clone PSR6561, mRNA sequence.
ACCESSION BE427000
VERSION BE427000.1 GI:9424933
KEYWORDS EST
SOURCE EST
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 640)
AUTHORS Anderson O.A., Appels R., Bailey P., Blake T., Close T., Cloutier
S., Dubcovsky J., Feuillet C., Gale M., Graner A., Gustafson P.,
Herrmann R.G., Holton T., Jacquemont J.M., Jia J., Joudrier P.,
Langridge P., Lazo G.R., Lin J.J., McGuire P., Ogihara Y.,
Pechioni N., Quilley C., Schuch W., Selvaraj G., Shariflou M.,
Sorrells M., Warburton M. and Wenzel G.
TITLE International Triticaceae EST Cooperative (ITPC): Production of
Expressed Sequence Tags for Species of the Triticeae
JOURNAL Unpublished (2000)
COMMENT Contact: Bailey P.
Cereals Group, John Innes Centre
Norwich, Norfolk NR4 7UH UNITED KINGDOM
Tel: 44 1603 452571 ext. 2587
Fax: 44 1603 502241
Email: paul.bailey@jic.ac.uk
International Triticaceae EST Cooperative (ITPC)

[illegible]

Accession	Definition	LOCUS	Result 13
CNS02EOD	Tetradon nigriviridis genome survey sequence. 262H4 of library G from Tetradon nigriviridis, genomic survey sequence.	681 bp	DNA linear GSS 13-MAY-2000

ACCESSION	AL193990
VERSION	AL193990.1
KEYWORDS	GSS: genome survey sequence
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis

REFERENCE	AUTHORS
1 (bases 1 to 681)	
Röest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,	

TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i>
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 681)
AUTHORS	Kost-Collins, R., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,

TITLE	JOURNAL	REFERENCE	NUMBERS
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	Unpublished	3 (bases 1 to 681)	Genome

REMARKS
JOURNAL
Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the *Tetradodon nigroviridis*
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/tetraodon>.

FEATURES	source
Location/Qualifiers	1. 681
/organism="Nectrodon nigroviridis"	
/db_xref="taxon:99883"	
/clone="362H14"	
/clone_11b="G"	
/note="Genoscope sequence ID : C0A262D0C7UPI-end : T7"	
BASE COUNT	74 a 105 c 170 g 282 t 50 others
ORIGIN	

Query Match 15.0%; Score 74.8; DB 12; Length 681;

Best Local Similarity 45.4%; Pred: No. 0.13;
Matches 173; Conservative 21; Mismatches 187; Indels 0; Gaps 0;

[illegible]

RESULT 14	
CNS03P82	
LOCUS	915 bp DNA linear
DEFINITION	Tetradon nigroviridis genome survey sequence PUC-ori end of clone
	DAF15 of library G from Tetradon nigroviridis, genomic survey
	sequence

ACCESSION	AL254315
VERSION	AL254315.1
KEYWORDS	GI:7975327
SOURCE	GSS: genome survey sequence.
ORGANISM	Tetraodon nigroviridis.
	Tetraodon nigroviridis.

REFERENCE
1 (bases 1 to 915) *Yersinia enterocolitica* strain 480/89, GenBank accession number G01001

TITLE	JOURNAL	REFERENCE
Characterization and repeat analysis of the compact genome of the freshwater putterfish <i>Tetraodon nigroviridis</i>	unpublished	2 (bases 1 to 915)

TITLE	Saurin, W. and Weissenbach, J.
JOURNAL	Human gene number estimate provided by genome wide analysis using
REFERENCE	Tetradodon nigroviridis DNA sequence
	Unpublished
	3 (bases 1 to 915)

AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigriviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .
FEATURES	Location/Qualifiers
source	1..915

```
/organism="Tetraodon nigroviridis"  
/db_xref="taxon:99883"
```

BASE COUNT	317 a	245 c	208 g	140 t	5 others
ORIGIN	/clone="04AF19" /clone_lib="G" /note="Genoscope sequence ID : C0B044CC10SF1-end : PUC-ori"				

Query Match	14.88;	Score 73.8;	DB 12;	length 915;
Best Local Similarity	54.98;	Pred. No. 0.18;		
Matches 168;	Conservative	0;	Mismatches 133;	Indels 5;
				Gaps 1;

QY	194	CAATCGCAATATGAGGACGCACCCAGGAGCATGCCAATGTGGAGAGCCCTTCACGGCTTTC	253
Db	394	CAATCGCAATGAGGACGCACCCAGGAGCATGCCAATGTGGAGAGCCCTTCACGGCTTTC	453
QY	254	CAATCGGAGGAGCGCTCTCGCTGAGCTTGCCACAGGAGGAGCGCCGCCACAGGCTTCGCCACGG	313
Db	454	CAATCGGAGGAGCGCTCTCGCTGAGCTTGCCACAGGAGGAGCGCCGCCACAGGCTTCGCCACGG	513
QY	314	GGAGCGCAACGAAGGGCATTTGGGCAACGAGGTGCATCTCCAGCGATCTGCTTACGAGAGGCG	373
Db	514	GGAGCGCAACGAAGGGCATTTGGGCAACGAGGTGCATCTCCAGCGATCTGCTTACGAGAGGCG	568
QY	374	CGTCCAGGGCATTCCTTAACGAGGTGGTGGCCGCCAGGCGCATGCGCCACAGAGAGCGCCGCCCA	433
Db	569	CGTCCAGGGCATTCCTTAACGAGGTGGTGGCCGCCAGGCGCATGCGCCACAGAGAGCGCCGCCCA	628
QY	434	GGAATTCGGCCAGGATGTGCGACAGGGCATGCCACAGGAGAGCGCCGCCAGGAGGATCGC	493
Db	629	GGAATTCGGCCAGGATGTGCGACAGGGCATGCCACAGGAGAGCGCCGCCAGGAGGATCGC	688
QY	494	CAACAA 499	
Db	689	CAACAA 694	

RESULT 15				
AV438522/c				
LOCUS	401 bp	mrna	linear	EST 23-AUG-2000
DEFINITION	Porphyra yezoensis 'YU-1'	Porphyra yezoensis	CDNA clone	
	PS037d03_L-5	mrna sequence.		

VERSION	AVA38522.1	GI:8593747
KEYWORDS	EST.	
SOURCE	Porphyra yezoensis.	

Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.

TITLE Tabata, S.
Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis*
JOURNAL DNA RES. 7, 223-227 (2000)

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
2-1-1 Higashi
Kazusa, Chiba 259-0292, Japan
Tel: 0438-780000 Fax: 0438-780001
E-mail: erika@kri.ac.jp

```

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
FEATURES
  Location/Qualifiers
    1..401

```

BASE COUNT	ORIGIN
37 a	187 c 124 g 53 t

Query Match 14.78; Score 73.4; DB 9; Length 401;

Best Local Similarity 51.8%; Pred. No. 0.22;
Matches 192; Conservative 0; Mismatches 176; Indels 3; Gaps 1.

Oy	122	CGAGGACCCATCCAGGAGCATGCCCAACGAGGAGGTGGCCAGGGCATCCCAATAGGGGT	181
Db	372	CGCGGGCGGGGGGTGGGCGAACCCGTCGGGTGTGTGTGGCAACCGGCGTTACGCGAGCCGT	313
Oy	182	CGCCGACAGGGCATGCGCAATGAGAGACGCCAACCCAGGGCATGCCAACTGGGAGCCGT	243
Db	312	CGCCGGCGCGGGGTGTAGGAGAGCGCATGTGGGAGATGACGGCACGGGCGTGGCGACCGGT	253
Oy	242	CCACGGCTTGGCCCAACGG---GGAGGCGTCTCTCAAGCTTGGCCAAACGGGACCGGCCCA	298
Db	252	GGGCGCTGATGGGACACGGGCTGGGCGACGCGTAGCGGGGGCGGGCGTCGGGACGGGT	193
Oy	299	GGGCGATCGCCACGGGGGAGCGCCACCAAGGGCATGGGCAACGAGGTCTACCATCCAGGGCAT	358
Db	192	GGGCGATGACGGCACGGGAGTGTGGGTGACGCCATGGGCAACGTACCGGACCGCGTGGGCGA	133
Oy	359	CGCTTACGAGGAGCGCGCTCCAGGGCATTCGTCTAAGAGGTGGCCGCCAGGGCATTCGCCAA	418
Db	132	CGCGGTCCCGGGCGCGGGTGGGAGCGCGGTGGGAGTACGGCACCGGCGCTCGGCGA	73
Oy	419	CGAGGACCCGCCAGGAGATGCCGAGAGATGTCTGACACGGGCACTCCCAACGAGAGACG	478
Db	72	CGCGGTGGCGCATATGAGGACACGGGGGTAGCGAGACCGGTAGCGGGGCGCGTCGCGGA	13
Oy	479	CGCCGACGGCA	489
Db	12	CGCGGTGGGAA	2

Search completed: October 10, 2002, 20:28:30
Job time : 698.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 84.9333 Seconds
(without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101a-16_COPY_1_500

Perfect score: 500
Sequence: 1 gagctacgcatctgtag.....ccacggatcgcacacaag 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457241 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Genesec_032802:*
2: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1980.DAT:*
3: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1981.DAT:*
4: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1982.DAT:*
5: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1983.DAT:*
6: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1984.DAT:*
7: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1985.DAT:*
8: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1986.DAT:*
9: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1987.DAT:*
10: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1988.DAT:*
11: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1989.DAT:*
12: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1990.DAT:*
13: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1991.DAT:*
14: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1992.DAT:*
15: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1993.DAT:*
16: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1994.DAT:*
17: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1995.DAT:*
18: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1996.DAT:*
19: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1997.DAT:*
20: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1998.DAT:*
21: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA1999.DAT:*
22: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA2000.DAT:*
23: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA2001A.DAT:*
24: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA2001B.DAT:*
25: /SIDSL/gcgdata/hold-genesec/genesecq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	748	22	AAD11123 Human small cell 1
2	500	100.0	1201	22	AAD11122 Human small cell 1
3	307.8	61.6	1593	23	AA572677 DNA encoding novel
4	302.4	60.5	3093	23	AA572678 DNA encoding novel
5	269.6	53.9	2922	23	AA572679 DNA encoding novel
6	257.6	51.5	518	22	ABA66861 Human foetal liver
7	257.6	51.5	518	22	ABA66861 Human foetal liver
8	257.6	51.5	518	22	AAK15297 Human brain expres
9	257.6	51.5	518	22	AAK15297 Human brain expres
					Probe #15758 used

10	250	50.0	3345	23	AA572676 DNA encoding novel
11	250	50.0	3522	23	AA572321 Human foetal liver
12	235	47.0	495	22	ABA54299 Probe #2522 for ge
13	235	47.0	495	22	ABA24056 Human brain expres
14	235	47.0	495	22	AAK02583 Probe #2635 used t
15	235	47.0	495	22	AAK33949 Human secreted pro
16	181.4	36.3	1530	21	AA116669 Human brain expres
17	178.2	35.6	537	22	AAK22452 Human bone marrow
18	178.2	35.6	537	22	AAK48620 Human bone marrow
19	178.2	35.6	537	22	AA124443 Probe #23134 used
20	178.2	35.6	537	22	AA154448 Human breast cell
21	178.2	35.6	576	22	ABA50147 Probe #12924 for g
22	178.2	35.6	576	22	AA122991 Human bone marrow
23	178.2	35.6	593	22	AAK09601 Human bone marrow
24	178.2	35.6	593	22	AAK35695 Human bone marrow
25	178.2	35.6	593	22	AA141410 Human foetal liver
26	171	34.2	586	22	ABA60111 Human brain expres
27	171	34.2	586	22	AAK08387 Human bone marrow
28	171	34.2	586	22	AAK34267 Human bone marrow
29	171	34.2	586	22	AA139989 Probe #675 used t
30	169.2	33.8	468	22	ABA25584 Probe #5050 for ge
31	169.2	33.8	468	22	AA115206 Human foetal liver
32	153.6	30.7	471	22	ABA72651 Human brain expres
33	153.6	30.7	471	22	AAK21080 Human bone marrow
34	153.6	30.7	471	22	AAK47234 Probe #21755 used
35	153.6	30.7	471	22	AA153069 DNA encoding novel
36	145.4	29.1	3127	23	ABA6376 Probe #14922 for g
37	142.4	28.5	471	22	ABA6456 Human breast cell
38	142.4	28.5	471	22	AA124385 Probe #14318 for g
39	140.8	28.2	459	22	ABA45009 Probe #1704 for ge
40	140.8	28.2	459	22	AA113771 Human foetal liver
41	140.8	28.2	526	22	ABA60045 Human brain expres
42	140.8	28.2	526	22	AAK08319 Human bone marrow
43	140.8	28.2	526	22	AAK34397 Probe #606 used t
44	140.8	28.2	526	22	AA139920 Human foetal liver
45	136	27.2	526	22	ABA61696

ALIGNMENTS

RESULT 1	AAD11123	standard; DNA: 748 BP.
ID	AAD11123	
XX		
AC	AAD11123;	
XX		
DT	24-SEP-2001 (first entry)	
XX		
DE	Human small cell lung cancer associated gene, DKFZp434C196 #2.	
XX		
KW	Human; small cell lung cancer; therapy; hCAP; nucleic acid; NM;	
KW	melanoma; cancer; colon; breast; head; neck; transitional cancer;	
KW	leiomyosarcoma; synovial sarcoma; cytostatic; ds.	
OS	Homo sapiens.	
XX		
PN	WO200153349-A2.	
XX		
PD	26-JUL-2001.	
XX		
PE	19-JAN-2001: 2001WO-US02015.	
XX		
PR	21-JUN-2000: 2000US-0489101.	
XX		
PA	(LUDW-) LUDWIG INST CANCER RES.	
PA	(SLOK) SLOAN KETTERING INST CANCER RES.	
PA	(CORR) CORNELL RES FOUND INC.	
XX		
PI	Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;	
XX	WPI; 2001-457597/49.	
DR		
XX		

PT Isolated polypeptide, used to treat or prognose a disorder
 PT characterized by expression of a hCAP e.g. cancer, is encoded by an
 PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
 XX
 PS Claim 57; Page 116; 152pp; English.

CC The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterised by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterised by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterised by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are small and non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
 CC The present sequence is human DKFZP434C196 DNA. This small cell lung
 CC cancer associated gene is designated as NY-SCLC-13.
 XX
 SO Sequence 748 BP; 166 A; 261 C; 246 G; 72 T; 3 other;

Query Match 100.0%; Score 500; DB 22; Length 748;
 Best Local Similarity 100.0%; Pred. No. 1.4e-67;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTCTACGGCATTCGCTGAGAGCGCTGCCAAGGCGATCGCTAATGAGAGCGCGACAGG 60
 DB 1 GAGTCTACGGCATTCGCTGAGAGCGCTGCCAAGGCGATCGCTAATGAGAGCGCGACAGG 60
 QY 61 GCATCGCTATGAGAGACACACAGTGCATCGCCCAAGAGAGCGCGCCAGGCGATCG 120
 DB 61 GCATCGCTATGAGAGACACACAGTGCATCGCCCAAGAGAGCGCGCCAGGCGATCG 120
 QY 121 CCGAGAGCGCATCCCGAGGCGATCGCCCAAGAGAGGTTGCCCGAGGCGATCGCCCAATGGGG 180
 DB 121 CCGAGAGCGCATCCCGAGGCGATCGCCCAAGAGAGGTTGCCCGAGGCGATCGCCCAATGGGG 180
 QY 181 TCGCCCGCAAGAGGCGATCGCCCAATGAGAGCGCGACACCGAGGCGATCGCCCAATGGGG 240
 DB 181 TCGCCCGCAAGAGGCGATCGCCCAATGAGAGCGCGACACCGAGGCGATCGCCCAATGGGG 240
 QY 241 TCCAGGCGCTTCGCGCAAGGCGCGCTCTCTAGCTTCGCGCAAGGCGCGCGCCAGG 300
 DB 241 TCCAGGCGCTTCGCGCAAGGCGCGCTCTCTAGCTTCGCGCAAGGCGCGCGCCAGG 300
 QY 301 GCATCGCCCAAGGCGCGATCGCCCAAGAGGCGATCGCCCAAGAGGCGATCG 360
 DB 301 GCATCGCCCAAGGCGCGATCGCCCAAGAGGCGATCGCCCAAGAGGCGATCG 360
 QY 361 CTAACGAGAGCGCGCTTCGCGCAAGGCGCGATCGCCCAAGAGGCGATCG 420
 DB 361 CTAACGAGAGCGCGCTTCGCGCAAGGCGCGATCGCCCAAGAGGCGATCG 420
 QY 421 AGAGCGCGCGCCCAAGGAGTTCGCGCAAGGCGCGATCGCCCAAGAGGCGCG 480
 DB 421 AGAGCGCGCGCCCAAGGAGTTCGCGCAAGGCGCGATCGCCCAAGAGGCGCG 480
 QY 481 CCCAGGCGATCGCCCAAGG 500
 DB 481 CCCAGGCGATCGCCCAAGG 500

RESULT 2
 AAD11122
 ID AAD11122 standard; DNA; 1201 BP.

AC AAD11122;
 XX 24-SEP-2001 (first entry)
 DE Human small cell lung cancer associated gene, DKFZP434C196 #1.
 XX

KM Human; small cell lung cancer; therapy: hCAP; nucleic acid; NA;
 KM melanoma; cancer; colon; breast; head; neck; transitional cancer;
 KM leiomyosarcoma; synovial sarcoma; cytostatic; ds.
 XX
 OS Homo sapiens.

XX WO200153349-A2.
 XX 26-JUL-2001.
 XX 19-JAN-2001; 2001WO-US02015.
 XX 21-JAN-2000; 2000US-0489101.
 XX (LUDWIG) LUDWIG INST CANCER RES.
 XX (SLOK) SLOAN KETTERING INST CANCER RES.
 XX (CORR) CORNELL RES FOUND INC.
 XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
 DR WPI; 2001-457597/49.

PT Isolated polypeptide, used to treat or prognose a disorder
 PT characterized by expression of a hCAP e.g. cancer, is encoded by an
 PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
 XX
 PS Claim 57; Page 114-115; 152pp; English.

CC The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterised by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterised by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterised by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are small and non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
 CC The present sequence is human DKFZP434C196 DNA. This small cell lung
 CC cancer associated gene is designated as NY-SCLC-13.
 XX

Sequence 1201 BP; 255 A; 377 C; 349 G; 178 T; 42 other;

Query Match 100.0%; Score 500; DB 22; Length 1201;
 Best Local Similarity 100.0%; Pred. No. 1.3e-67;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTCTACGGCATTCGCTGAGAGCGCTGCCAAGGCGATCGCTAATGAGAGCGCGACAGG 60
 DB 1 GAGTCTACGGCATTCGCTGAGAGCGCTGCCAAGGCGATCGCTAATGAGAGCGCGACAGG 60
 QY 61 GCATCGCTATGAGAGACACACAGTGCATCGCCCAAGAGAGCGCGCCAGGCGATCG 120
 DB 61 GCATCGCTATGAGAGACACACAGTGCATCGCCCAAGAGAGCGCGCCAGGCGATCG 120
 QY 121 CCGAGAGCGCATCCCGAGGCGATCGCCCAAGAGAGGTTGCCCGAGGCGATCGCCCAATGGGG 180
 DB 121 CCGAGAGCGCATCCCGAGGCGATCGCCCAAGAGAGGTTGCCCGAGGCGATCGCCCAATGGGG 180
 QY 181 TCGCCCGCAAGAGGCGATCGCCCAATGAGAGCGCGACACCGAGGCGATCGCCCAATGGGG 240
 DB 181 TCGCCCGCAAGAGGCGATCGCCCAATGAGAGCGCGACACCGAGGCGATCGCCCAATGGGG 240
 QY 241 TCCAGGCGCTTCGCGCAAGGCGCGCTCTCTAGCTTCGCGCAAGGCGCGCGCCAGG 300
 DB 241 TCCAGGCGCTTCGCGCAAGGCGCGCTCTCTAGCTTCGCGCAAGGCGCGCGCCAGG 300
 QY 301 GCATCGCCCAAGGCGCGATCGCCCAAGAGGCGATCGCCCAAGAGGCGATCG 360
 DB 301 GCATCGCCCAAGGCGCGATCGCCCAAGAGGCGATCGCCCAAGAGGCGATCG 360
 QY 361 CTAACGAGAGCGCGCTTCGCGCAAGGCGCGATCGCCCAAGAGGCGATCG 420

DB 361 CTACACGACGACGCGCCGACGACGATCGCTAACGAGTGGCGCCGACGACGATCGCCCAACG 420
 QY 421 AGACGCGCGCGCGACGAGATCGCCGACGATGTCGACACGAGGCGATCGCCCAACGAGCGCCG 480
 DB 421 AGACGCGCGCGCGACGAGATCGCCGACGATGTCGACACGAGGCGATCGCCCAACGAGCGCCG 480
 QY 481 CCCAGGCGATCGCCCAACG 500
 DB 481 CCCAGGCGATCGCCCAACG 500

RESULT 3

AAS72677
 ID AAS72677 standard; cDNA, 1593 BP.

AC AAS72677;

BT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #8481.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB: ABG08490.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 1: SEQ ID No 8481; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS6197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WFO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 1593 BP; 444 A; 429 C; 434 G; 286 T; 0 other;

Query Match 61.6%; Score 307.8; DB 23; Length 1593;
 Best Local Similarity 78.8%; Pred. No. 1,2e-38;
 Matches 393; Conservative 0; Mismatches 102; Indels 4; Gaps 2;

QY 5 CTACGCGATGCTGAGACGCTCCCGACGAGATGCTATATGAGAGACCGACGACGAT 64
 DB 873 CGACATCGCTAACGAGATGCTGTCACGAGATGCTAACGAGAGCGCGACGACGAT 932
 QY 65 CGTATATGAGACACACCCAGTGCATCGCCCAACGAGAGACCGCCAGGATCGC--- 121
 DB 933 CGCCACGAGAGAGCGCCGACGAGATCGCCCAACGAGAGACCGCCAGGATCGCAG 992
 QY 122 CGAGAGCGCATCGCCCAACGAGAGATGCTTCCGACGAGATCGCCCAACGAGAGAT 181
 DB 993 CGAGAGATCGCCGACGAGATCGCCCAACGAGAGATGCTTCCGACGAGATCGCCCAACGAGAG 1052
 QY 182 CGCGGACAGAGGATGCTGCTATGAGAGAGCGACGAGAGATCGCCCAACGAGAGAT 241
 DB 1053 CGCGGACAGAGGATGCTGCTATGAGAGAGCGACGAGAGATCGCCCAACGAGAGAT 1112
 QY 242 CCACGCTTCGCCACAGGAGCGCGCTCTCTAGCTTCCGACGAGAGATCGCCCAACGAGAG 301
 DB 1113 CCACGAGATCGCCCAACGAGAGAGCGCGCGCGAGGATGCTTCCGACGAGAGATCGCCCAACGAGAG 1172
 QY 302 CATCGCCCAACGAGAGAGCGCGCGCGAGGATGCTTCCGACGAGAGATCGCCCAACGAGAGAT 361
 DB 1173 CATCGCCCAACGAGAGAGCGCGCGCGAGGATGCTTCCGACGAGAGATCGCCCAACGAGAGAT 1232
 QY 362 TACGAGAGAGCGCGCGCGAGGATGCTTACGAGAGATGCTTCCGACGAGAGATCGCCCAACGAGAG 421
 DB 1233 TACGAGAGATGCTTCCGACGAGATGCTTACGAGAGATGCTTCCGACGAGAGATGCTTACGAGAGAT 1292
 QY 422 GAGAGCGCGCGCGAGGATGCTTCCGACGAGAGATGCTTCCGACGAGAGATGCTTCCGACGAGAGAT 481
 DB 1293 GAGAGCGCGCGCGAGGATGCTTACGAGAGATGCTTCCGACGAGAGATGCTTCCGACGAGAGAT 1351
 QY 482 CCAGGCGATCGCCCAACG 500
 DB 1352 ACAGGATGCTTACGAG 1370

RESULT 4

AAS72678
 ID AAS72678 standard; cDNA; 3093 BP.

AC AAS72678;

BT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #8482.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB: ABG08491.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 1: SEQ ID No 8482; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations in
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 3093 BP; 875 A; 904 C; 862 G; 452 T; 0 other;

Query Match 60.5%; Score 302.4; DB 23; Length 3093;
Best Local Similarity 79.8%; Pred. No. 7; 2e-38;
Matches 389; Conservative 0; Mismatches 91; Indels 6; Gaps 2;

24 GCTGCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 83
1990 GATGCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2049
84 CAGTGCATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 140
2050 CAGTGCATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2109
141 ATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 200
2110 ATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2169
201 AATGAGAGCGCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 260
2170 AATGAGAGCGCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2229
261 GAGCGCGCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 320
2230 GAGCGCGCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2289
321 AATGAGAGCGCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 380
2290 AATGAGAGCGCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2349
381 GCGATATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 440
2350 GCGATATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2409
441 GC---CGAGAGATGTCGACAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 497
2410 GCGAGAGATGTCGACAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2469

RESULT 5
AAS72679
ID AAS72679 standard; cDNA; 2922 BP.
XX
AC AAS72679;

XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #8483.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensics;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HSE-) HXSE0 INC.
XX Dymnac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG08492.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 1: SEQ ID No 8483; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations in
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 2922 BP; 833 A; 829 C; 806 G; 454 T; 0 other;

Query Match 53.9%; Score 269.6; DB 23; Length 2922;
Best Local Similarity 75.4%; Pred. No. 6; 3e-33;
Matches 362; Conservative 0; Mismatches 114; Indels 4; Gaps 2;

24 GCTGCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 83
1906 GATGCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 1965
84 CAGTGCATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 140
1966 CAGTGCATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2025
141 ATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 200
2025 ATCCCAAGGCGATCCCTAATAGAGACGCCCAAGGCGATCGTATAGAGACACCAC 2085

100

Penn SG, Hanzel DK, Chen W, Rank DR,

XX WPI: 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PS Claim 4: SEQ ID NO 12396; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 518 BP; 140 A; 170 C; 139 G; 69 T; 0 other;

Query Match 51.5%; Score 257.6; DB 22; Length 518;
 Best Local Similarity 75.2%; Pred. No. 5.1e-31;
 Matches 376; Conservative 0; Mismatches 114; Indels 10; Gaps 4;

QY 5 CTACGGCATTGCTGAGGAGCGCTGCCAGGCGATCGCTAATGAGAGCGCCGACGAGGCGAT 64
 DB 23 CACATCACTACTACGAGAGTGCCTCAGACGATATCATACGAGAGCGCTTCCACGCGAT 82
 QY 65 CGCTAATGAGAGACACCACCACTGATCGCCACGAGAGACGCCCGAGGCGATCGC--- 121
 DB 83 CGCTAATGAGAGACGCCCGAGGCGATCGCCACGAGAGACGCCCGAGGCGATCGC--- 142
 QY 122 CGAGGAGCGCATTCAGAGGCGATCGCCACGAGAGAGTGGCCAGGCGATCGCCAAATGGGGT 181
 DB 143 CGAGGAGCGCGCCGACGAGATCGCCACGAGAGAGCGCCCGAGGCGATCGCCAAATGGGGT 202
 QY 182 CGCGGAGAGGCGATCGCCAAATGAGAGAGCGCCAGGCGATCGCCAAATGGGGT 241
 DB 203 CGCGGAGAGGCGATCGCCAAATGAGAGAGCGCCCGAGGCGATCGCCAAATGGGGT 262
 QY 242 CGAGGCGCTTCGCAACGGGGAGCGCGCTTCGATCGCCACGAGAGAGCGCCCGAGG 301
 DB 263 CGATGCGATCGCCACGAGAGAGCGCGCTTCGATCGCCACGAGAGAGCGCGCTTCG 322
 QY 302 CATCGGCAACGGGAGCGCGACCAAGGCGATCGCCACGAGAGAGCGCGCTTCG 361
 DB 323 CATCGGTAATGAGAGACTCGCTATAGACATCGCTAATGAGAGAGCGCTTCG 382
 QY 362 TAACGAGAGCGCGCTTCGAGGCGATCGCCAAATGAGAGAGCGCGCGCGCGCGCGCG 421
 DB 383 TA---ATGACACCGTACAGAGGCA-CGCTTACGAGAGAGTGTACACGACGCTGCTATATA 438
 QY 422 GAGAGCGCGCGCGAGGAGATCG---CGAGAGATGTCACAGAGGCGATCGCCACGAGAGCGC 478
 DB 439 GGCACACCGTACAGAGGCGATCGCTAATGAGAGAGTGTATATAGATCGCTTACACAGAGAC 498
 QY 479 CGCGGAGGCGATCGCCACGA 498
 DB 499 TCTACAGAGCGCTGCTTACA 518

RESULT 8
 AAK15297
 ID AAK15297 standard; DNA; 518 BP.
 XX
 AC AAK15297;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 15288.

XX Human: brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-0500667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207436.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PS Example 4: SEQ ID NO: 15288; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 518 BP; 140 A; 170 C; 139 G; 69 T; 0 other;

Query Match 51.5%; Score 257.6; DB 22; Length 518;
 Best Local Similarity 75.2%; Pred. No. 5.1e-31;
 Matches 376; Conservative 0; Mismatches 114; Indels 10; Gaps 4;

QY 5 CTACGGCATTGCTGAGGAGCGCTGCCAGGCGATCGCTAATGAGAGCGCCGACGAGGCGAT 64
 DB 23 CACATCACTACTACGAGAGTGCCTCAGACGATATCATACGAGAGCGCTTCCACGCGAT 82
 QY 65 CGCTAATGAGAGACACCACCACTGATCGCCACGAGAGAGCGCCCGAGGCGATCGC--- 121
 DB 83 CGCTAATGAGAGACGCCCGAGGCGATCGCCACGAGAGAGCGCCCGAGGCGATCGC--- 142
 QY 122 CGAGGAGCGCATTCAGAGGCGATCGCCACGAGAGAGTGGCCAGGCGATCGCCAAATGGGGT 181
 DB 143 CGAGGAGCGCGCCGACGAGATCGCCACGAGAGAGCGCCCGAGGCGATCGCCAAATGGGGT 202
 QY 182 CGCGGAGAGGCGATCGCCAAATGAGAGAGCGCCCGAGGCGATCGCCAAATGGGGT 241
 DB 203 CGCGGAGAGGCGATCGCCAAATGAGAGAGCGCCCGAGGCGATCGCCAAATGGGGT 262
 QY 242 CGAGGCGCTTCGCAACGGGGAGCGCGCTTCGATCGCCACGAGAGAGCGCCCGAGG 301
 DB 263 CGATGCGATCGCCACGAGAGAGCGCGCTTCGATCGCCACGAGAGAGCGCGCTTCG 322
 QY 302 CATCGGCAACGGGAGCGCGACCAAGGCGATCGCCACGAGAGAGCGCGCTTCG 361
 DB 323 CATCGGTAATGAGAGACTCGCTATAGACATCGCTAATGAGAGAGCGCTTCG 382
 QY 362 TAACGAGAGCGCGCTTCGAGGCGATCGCCAAATGAGAGAGCGCGCGCGCGCGCGCG 421
 DB 383 TA---ATGACACCGTACAGAGGCA-CGCTTACGAGAGAGTGTACACGACGCTGCTAATGA 438

QY 422 GGAGCGCCCGCCAGGATGCG--CCGAGATGTCGACAGGCGATCGCCACGAGAGCG 478
 DB 439 GGACCCCATATCAGAGCGATCGGTAATGAGAGAGCGTGTATATGATCATCGCTAACAGAGACAC 498
 QY 479 CGCCAGGCGATCGCCACCA 498
 DB 499 TCTACAGCGCTCGCTACCA 518

RESULT 9

AA147072
 ID AA147072 standard; DNA; 518 BP.

AA147072:

17-OCT-2001 (first entry)

DE Probe #15758 used to measure gene expression in human placenta sample.

KM Probe: microarray; human; placenta; antenatal diagnosis;

KM genetic disorder; ss.

OS Homo sapiens.

PM WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000US-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

PS Claim 25; SEQ ID No 15758; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.

XX Sequence 518 BP; 140 A; 170 C; 139 G; 69 T; 0 other;

QY Query Match 51.5%; Score 257.6; DB 22; Length 518;

DB Best Local Similarity 75.2%; Pred. No. 5.1e-31;

Matches 376; Conservative 0; Mismatches 114; Indels 10; Gaps 4;

QY 182 CGCCGACAGGGCATCGCCAAATGAGAGAGCGCCACCCAGGCGATCGCCACGCGCGT 241
 DB 203 CGCGCGCCGCGATCGCCAAATGAGAGAGCGCGCCACCCAGGCGATCGCCACGAGAGCGCGT 262
 QY 242 CCAAGGCTTGGCCAAAGGAGAGCGCGTCTTACGTTGCGCAACGAGGAGCGCGCCAGG 301
 DB 263 CCAAGGCGATCGCCAAAGGAGAGCGCGTCTTACGTTGCGCAACGAGGAGCGCGCGTCAAG 322
 QY 302 CATCGCCAAAGGAGAGCGCGCCAAAGGAGAGCGCGTCTTACGTTGCGCAACGAGGAGCG 361
 DB 323 CATCGCAATAGAGAGAGCGCGTCTTACGTTGCGCAACGAGGAGAGCGCGTCTTACGATCGC 382
 QY 362 TAAAGAGAGCGCGTCTTACGTTGCGCAACGAGGAGAGCGCGCGCCAGGCGATCGCCAAAG 421
 DB 383 TA---ATGACAGCGTACAGAGCA--CGCTAACAGAGAGTGTGTACAGCATCGCTAATGA 438
 QY 422 GGAGCGCCCGCCAGGATGCG--CCGAGATGTCGACAGGCGATCGCCACGAGAGAGCG 478
 DB 439 GGACCCCATATCAGAGCGATCGGTAATGAGAGAGCGTGTATATGATCATCGCTAACAGAGAC 498
 QY 479 CGCCAGGCGATCGCCACCA 498
 DB 499 TCTACAGCGCTCGCTACCA 518

RESULT 10

AA572676
 ID AA572676 standard; cDNA; 3345 BP.

AA572676:

13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #8480.

KM Human: chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PM WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG08489.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

PS Claim 1; SEQ ID No 8480; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC for restoring normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pcl_sequences.

XX Sequence 3345 BP; 954 A; 940 C; 913 G; 538 T; 0 other;

Query Match 50.0%; Score 250; DB 23; Length 3345;
 Best Local Similarity 74.1%; Pred. No. 5,6e-30;
 Matches 372; Conservative 0; Mismatches 120; Indels 10; Gaps 4;

OY 5 CTACGGCATTCGTCGAGGAGCGCTCCAGGATCGCTAATGAGAGCGCCGACCGGCGCAT 64
 DB 2076 CAACATCTACTACGAGATGCTCTACAGATATCTACTACGAGAGCGCTGTCCAGCGCAT 2135
 OY 65 CGCTAATGAGAGACCCACCCACTGCTATCCACAGAGAAAGCCGCCAGGCGATCGC--- 121
 DB 2136 CGCTAATGAGAGCGCCGACCAAGGATGCGCCACAGAGAGAGCGCCGCCAGGCGATCGC 2195
 OY 122 CGAGAGCGCCATCCAGGAGATGCGCCACAGAGAGTTGCCAGGCGATCGCCATATGGGT 181
 DB 2196 CGAGAGCGCCGCGCCGAGATGCGCCAGAGAGCGCCGCCAGGCGATCGCCACAGCGAT 2255
 OY 182 CGCCGACAGGAGGATCGCCATGAGAGAGCGCCAGGCGATCGCCATGAGGAGCGCGT 241
 DB 2256 CGCCGCGCCAGGAGGATCGCCACAGAGAGCGCCGCCAGGCGATCGCCACAGAGCGCGC 2315
 OY 242 CGAGGAGGTCGCGCAAGGAGAGCGCGCTCTGCTTGGCCACAGGAGAGCGCCCGCCAGG 301
 DB 2316 CGAGGAGGTCGCGCAAGAGAGATGCTGCCAGGCGATGCTTAAGAGAGAGTCGGTGACGG 2375
 OY 302 CATCGGCAACGGGAGCGCCACCAAGGATGAGGAGAGAGGATCGCATCGGATCGCGC 361
 DB 2376 CATCGGTAATGAGAGACTCGTATACGATCGCTAATGAGAGAGTCGCTATATGATCGC 2435
 OY 362 TAACGAGAGAGCGCGCTCCAGGATCGCTAATGAGAGAGTGGCCGCCAGGATCGCCACGA 421
 DB 2436 TA---ATGACACCGTACAGAGCA-CACTAACGAGAGCGCTGTACAGACATCGCTAATGA 2491
 OY 422 GAGAGCGCCCGCCAGGAGATCG---CCGAGAGATGTCGACAGGAGATCGCCACAGAGAGCG 478
 DB 2492 GAGAGCGCCATACAGGATGCTAATGAGAGAGCGTTGTATATGATCGCTAACAGAGACAC 2551
 OY 479 CGCCGAGGCGATCGCCACAG 500
 DB 2552 TCTACAGAGCGCTCGCTAACAG 2573

RESULT 11
 AAS72321
 ID AAS72321 standard; cDNA; 3522 BP.
 XX AAS72321;
 XX 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #8125.
 XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 OS M0200175067-A2.
 XX

PD 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-0508631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG08134.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID No 8125; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pcl_sequences.

XX Sequence 3522 BP; 1005 A; 1006 C; 959 G; 552 T; 0 other;

Query Match 50.0%; Score 250; DB 23; Length 3522;
 Best Local Similarity 74.1%; Pred. No. 5,6e-30;
 Matches 372; Conservative 0; Mismatches 120; Indels 10; Gaps 4;

OY 5 CTACGGCATTCGTCGAGGAGCGCTCCAGGATCGCTAATGAGAGCGCCGACCGGCGCAT 64
 DB 2253 CAACATCTACTACGAGATGCTCTACAGATATCTACTACGAGAGCGCTGTCCAGCGCAT 2132
 OY 65 CGCTAATGAGAGACCCACCCACTGCTATCCACAGAGAAAGCCGCCAGGCGATCGC--- 121
 DB 2313 CGCTAATGAGAGCGCCGACCAAGGATGCGCCACAGAGAGAGCGCCGCCAGGCGATCGC 2372
 OY 122 CGAGAGCGCCATCCAGGAGATGCGCCACAGAGAGTGGCCGCCAGGATCGCCATATGGGT 181
 DB 2373 CGAGAGCGCCGCGCCAGGATGCGCCACAGAGAGCGCGCCGCCAGGCGATCGCCACAG 2432
 OY 182 CGCCGACAGGCGATCGCCATGAGAGAGCGCCACAGGAGATCGCCATATGGGATCGCGT 241
 DB 2433 CGCCGCGCCAGGAGGATCGCCACAGAGAGCGCCGCCAGGCGATCGCCACAGAGCGCGC 2492
 OY 242 CGAGGAGGTCGCGCAAGGAGAGCGCGCTCTGCTTGGCCACAGGAGAGCGCCCGCCAGG 301
 DB 2493 CGAGGAGGTCGCGCAAGGAGATGCTCTGCTTGGCCACAGGAGAGCGCGCGTGACGG 2552
 OY 302 CATCGGCAACGGGAGCGCCACCAAGGATGAGGAGAGAGGATCGCATCGGATCGCGC 361
 DB 2553 CATCGCTAATGAGAGACTCGTATACGATCGCTAATGAGAGAGTGGCATATATACATCGC 2612

QY 362 TAACGAGACGCCGTCACGAGGATCGCTACAGAGTGGCCGCCAGGCGCATCGCCACGA 421
 DB 2613 TA---ATGACACCGTACACAGGCA-CACATACGAGAGCGCTGTACACGACATCGCTATGCA 2668
 QY 422 GGAGCGCCGCCAGGGAATG---CCGAGATGTCGACAGGCGATCGCCACAGAGAGCGC 478
 DB 2669 GGACACCATACAGGATGTTATGAGAGCTGTATATGACATCGCTACAGAGACAC 2728
 QY 479 CGCCAGGCGATCGCCACAG 500
 DB 2729 TCTACAGCGCTCGCTACACAG 2750

RESULT 12
 ABA54299
 ID ABA54299 standard; DNA; 495 BP.
 AC ABA54299;
 DT 01-FEB-2002 (first entry)
 XX Human foetal liver single exon nucleic acid probe #2604.
 XX Human foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX Homo sapiens.
 XX WO20015727-A2.
 XX 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00669.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 PS Claim 1; SEQ ID NO 2604; 639bp + sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;

Query Match 47.0%; Score 235; DB 22; Length 495;
 Best local similarity 77.3%; Pred. No. 1.3e-27;
 Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

QY 43 ATGAGAGCGCGGATGCTATGAGAGACAGCAGCCATGCGTCCAGAGAGG 102
 DB 108 ATGAGAGACCGGTACACATACATACAGAGATGCTCCACAGATACATACAGAGG 167
 QY 103 AAGCCGCCAGGCGATCGC---CGAGAGCGCATCTCAGAGGATCGCCACAGAGAGGTTG 159

DB 168 ACGTGTCCACGAGCATCGCTTAAGAGGCGCGCGCCACGAGGATCGCCACAGAGAGCGCG 227
 QY 160 CCCAGGCGATCGCCATGAGGTCGCGCCGACAGGAGCATCGCCATGAGAGAGCCACCCAGG 219
 DB 228 CCCAGGCGATCGCCACAGAGAGCGCGCCGACAGGAAATCGCCACAGAGAGCGCGCGAGG 287
 QY 220 GCATCGCCAACTGAGAGCGCGCTCCAGGCTTCGCCAACGAGGAGCGCGCTTCAGCTTCG 279
 DB 288 GCATCGCCAACTGAGAGCGCGCGCCGACAGGAGCATCGCCACAGAGAGCGCGCGCGAGGATCG 347
 QY 280 CCAGGAGGAGCGCGCGCGGATCGCCACAGGAGCGCGCCACAGAGGCGATCGGCGACAG 339
 DB 348 CCAGGAGGAGCGCGCGCGGATCGCCACAGGAGCGCGCGCCACAGAGGCGATCGGCGACAG 407
 QY 340 AGCTCAGCATCGCCAGGATCGCTTAAGAGAGCGCGCTTCAGAGGATCGCTTAAGAGATG 399
 DB 408 AGGATCGCTGCGAGGATCGCTTAAGAGAGCGCGCTTCAGAGGATCGCTTAAGAGATG 467
 QY 400 CGCGCCAGGCGATCGCCACAGAGAGCG 426
 DB 468 CCATATATGACATCGCTTAAGAGAGCG 494

RESULT 13
 ABA24056
 ID ABA24056 standard; DNA; 495 BP.
 AC ABA24056;
 DT 23-JAN-2002 (first entry)
 XX Probe #2522 for gene expression analysis in human heart cell sample.
 XX Human; gene expression; heart; microarray; vascular system; probe;
 XX cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease; ss.
 XX Homo sapiens.
 XX WO20015727-A2.
 XX 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00666.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488899/53.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PS Claim 1; SEQ ID NO 2522; 530bp; English.
 XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging, the
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g., cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;
 SQ
 Query Match 47.0%; Score 235; DB 22; Length 495;
 Best Local Similarity 77.3%; Pred. No. 1.3e-27;
 Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;
 XX
 YY 43 ATGAGGAGCCGACGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 102
 DB 108 ATGAGGAGCCGACGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 167
 YY 103 AAGCCGCCAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 159
 DB 168 AAGCCGCCAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 227
 YY 160 CCGAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 219
 DB 228 CCGAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 287
 YY 220 GCATGCGCAAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 279
 DB 288 GCATGCGCAAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 347
 YY 280 CCAAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 339
 DB 348 CCAAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 407
 YY 340 AGGTACACATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 399
 DB 408 AGGTACACATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 467
 YY 400 CCGCCGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 426
 DB 468 CCGCCGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 494

RESULT 14
 ID AAK02583 standard; DNA; 495 BP.
 AC AAK02583;
 XX
 YY 05-NOV-2001 (first entry)
 DB Human brain expressed single exon probe SEQ ID NO: 2574.
 YY Human: brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 OS Homo sapiens.
 PN WO200157275-A2.
 PD 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 YY 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI

XX WPI; 2001-483446/52.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 XX brains -
 PT
 PS Example 4; SEQ ID NO: 2574; 650bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;
 XX
 Query Match 47.0%; Score 235; DB 22; Length 495;
 Best Local Similarity 77.3%; Pred. No. 1.3e-27;
 Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;
 XX
 YY 43 ATGAGGAGCCGACGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 102
 DB 108 ATGAGGAGCCGACGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 167
 YY 103 AAGCCGCCAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 159
 DB 168 AAGCCGCCAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 227
 YY 160 CCGAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 219
 DB 228 CCGAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 287
 YY 220 GCATGCGCAAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 279
 DB 288 GCATGCGCAAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 347
 YY 280 CCAAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 339
 DB 348 CCAAGGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 407
 YY 340 AGGTACACATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 399
 DB 408 AGGTACACATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 467
 YY 400 CCGCCGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 426
 DB 468 CCGCCGAGGATGCTATGAGACACACCCAGTGCATGCGCAAGAGG 494

RESULT 15
 ID AAI33949 standard; DNA; 495 BP.
 AC AAI33949;
 XX
 YY 17-OCT-2001 (first entry)
 DB Probe #2635 used to measure gene expression in human placenta sample.
 YY Probe: microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 PN WO200157272-A2.
 PD 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-488897/53.

XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -

PS C1a1m 25; SEQ ID NO 2635; 654bp; English.

XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.

XX
SQ Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;

Query Match 47.0%; Score 235; DB 22; Length 495;

Best Local Similarity 77.3%; Pred. No. 1,3e-27;

Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

QY 43 ATGAGGACCCGACGCGGATCGCTAATGAGACACCCAGTGCATGCCACAGAG 102
DB 108 ATGAGGACCCGATACACATGCTACTACAGAGAGTGCCTCAGACGATATCTACAGAG 167
QY 103 AAGCCGCCAGGCGATCGC---CGAGAGCGCCATCCAGGCGATGCCAACAGAGAGTTG 159
DB 168 ACGCTGTCCAGGGATCGCTAACGAGGCGCGCGCCAGGCGATGCCAACAGAGAGCGCG 227
QY 160 CCCAGGGCATCGCCATGAGGTCGCCGACACAGGGCATCGCCATGAGAGAGCGCCACCG 219
DB 228 CCCAGGGCATCGCCACGAGGAGGAGCGCGCCAGGATCGCCAGGAGAGAGCGCGCCAG 287
QY 220 GCATCGCACTGAGGAGCGCGTCCAGAGGCTTCCCAAGGAGGAGCGCGTCTGAGTTG 279
DB 288 GCATCGCCCAAGAGAGTCCGCGCGCCAGGCGATGCCAACAGAGAGCGCGCCAGGCGAT 347
QY 280 CCAAGGGGAGAGGCGCGCGCCAGGCGATGCCAACAGGAGGAGGAGGAGGAGGAGGAG 339
DB 348 CCAAGGAGAGAGGCGCGTCCATGCGATGCCAACAGAGAGAGCGTCCAGGCGATGCCAAC 407
QY 340 AGGTACCATCCAGGCGATGCCAACAGAGAGCGCGTCCAGGCGATGCCAACAGAGAGT 399
DB 408 AGGATCGCGTCCAGGCGATGCCAACAGAGAGCGTCCAGGCGATGCCAACAGAGAGT 467
QY 400 CCGCCCGAGGCGATGCCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
DB 468 CCATATATGAGATCGCTAATGAGAGCGG 494

Search completed: October 10, 2002, 15:18:05
Job time : 93 secs

Fri Oct 11 09:30:45 2002

us-09-489-101a-15_copy_1_500.rni

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 18.7333 Seconds
(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101A-15_COPY_1_500

Perfect score: 500
Sequence: 1 gagctacgcatctgtctgag.....ccagagcatcgcacacaaag 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116.4	23.3	390	4	US-09-197-649-7
2	87.4	17.5	4403765	4	US-09-103-840A-2
3	87.8	16.6	390	4	US-09-197-649-7
4	77.2	15.4	1140	3	US-09-023-173-4
5	77.7	15.4	1494	4	US-09-056-556-176
6	76.8	15.4	1931	2	US-09-130-114-2
7	76.2	15.4	1439	4	US-09-056-556-167
8	74.4	14.9	1248	4	US-09-105-357-1
9	74.4	14.9	1439	4	US-09-056-556-167
10	74.4	14.9	13613	4	US-09-105-357-3
11	74	14.8	5036	4	US-09-177-349-2
12	74	14.8	5970	3	US-08-320-878-21
13	73.4	14.7	12588	2	US-08-387-942C-1
14	73.2	14.6	15872	3	US-09-105-537-1
15	72.6	14.5	1182	4	US-09-025-691-2
16	72.4	14.5	1182	4	US-09-385-028-19
17	72.4	14.5	11604	4	US-09-385-028-13
18	72.4	14.5	15079	4	US-09-385-028-1
19	72.2	14.4	400	4	US-09-056-556-179
20	72.2	14.4	3306	1	US-08-261-206A-71
21	72	14.4	494	4	US-09-056-556-176
22	72	14.4	985	4	US-09-056-556-182
23	71.8	14.4	774	3	US-08-956-307B-11
24	71.8	14.4	778	3	US-08-956-307B-12
25	70.8	14.2	4411529	4	US-09-103-840A-1
26	70.8	14.2	2712	3	US-09-025-691-4
27	70.8	14.2	4403765	4	US-09-103-840A-2

28	70.8	14.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
C 29	70.2	14.0	30001	2	US-08-125-468-1	Sequence 1, Appli
C 30	70.2	14.0	30001	1	US-08-474-933-1	Sequence 1, Appli
C 31	69.2	13.8	985	4	US-09-056-556-182	Sequence 182, App
C 32	69	13.8	12588	2	US-08-387-942C-1	Sequence 23, Appli
C 33	67.6	13.5	1193	4	US-09-372-422A-23	Sequence 23, Appli
C 34	67	13.4	281	2	US-08-419-075-4	Sequence 25, Appli
C 35	66.8	13.4	352	2	US-08-419-075-25	Sequence 2, Appli
C 36	66.2	13.2	1187	1	US-08-440-856A-2	Sequence 17, Appli
C 37	66	13.2	579	2	US-08-737-524B-17	Sequence 12, Appli
C 38	66	13.2	579	2	US-08-419-075-3	Sequence 2, Appli
C 39	66	13.2	639	2	US-08-737-524B-12	Sequence 20, Appli
C 40	66	13.2	639	2	US-08-419-075-2	Sequence 2, Appli
C 41	66	13.2	647	2	US-08-419-075-20	Sequence 18, Appli
C 42	66	13.2	653	2	US-08-419-075-18	Sequence 9, Appli
C 43	66	13.2	2123	2	US-08-737-524B-9	Sequence 31, Appli
C 44	66	13.2	2123	2	US-08-419-075-1	Sequence 31, Appli
C 45	65.8	13.2	459	2	US-08-387-942C-31	

ALIGNMENTS

RESULT 1
US-09-197-649-7
Sequence 7, Application US/09197649
Patent No. 6194550

GENERAL INFORMATION:

APPLICANT: Tuery, Craig

APPLICANT: Smith, Jonathan D.

TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation

APPLICANT: Sutliff, Thomas D.
APPLICANT: Rodriguez, Raymond L.
TITLE OF INVENTION: Production of Mature Proteins
TITLE OF INVENTION: In Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Denlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023.173
FILING DATE: 13-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/038,168
FILING DATE: 13-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Pelthor, Joanne R
REGISTRATION NUMBER: P42995
REFERENCE/DOCKET NUMBER: 0665-0007.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: codon-optimized RAY3D signal-prosubtilisin BPN'
US-09-023-173-4

Query Match 15.4%; Score 77.2; DB 3; Length 1140;
Best Local Similarity 48.8%; Pred. No. 3.6e-06;
Matches 239; Conservative 0; Mismatches 248; Indels 3; Gaps 1;
QY 8 CGGATGTCGAGAGCGTCCCGGATGCTAATGAGAGCGCCGACGAGGCGATGCG 67
DB 270 CGCTAGCTGAG 329
QY 68 TATGAG 127
DB 330 CAGCGATGAG 389
QY 128 CGGATCCAG 187
DB 390 CGCGATGATGAG 449
QY 188 ACAGGAGATGAG 247
DB 450 TAGCATGATGAG 509
QY 248 CTTGCGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 307
DB 510 CGCGAG 569
QY 308 CAGCGAG 367
DB 570 CAGCGTCAAG 629
QY 368 GAGCGAG 424
DB 630 CAGCGAG 689

QY 425 CGCGCGAG 484
DB 690 CGCGAG 749
QY 485 GAGCAG 494
DB 750 CGTGTGCGAG 759

RESULT 5
US-09-056-556-176
Sequence 176, Application us/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelley, Yashir A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056.556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mark, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-056-556-176

Query Match 15.4%; Score 77; DB 4; Length 494;
Best Local Similarity 49.6%; Pred. No. 3.8e-06;
Matches 197; Conservative 0; Mismatches 200; Indels 0; Gaps 0;
QY 38 CGCTATGAG 97
DB 16 CGCGAG 75
QY 98 CGAG 157
DB 76 CCAAG 135
QY 158 TGCCAG 217
DB 136 CGCGAG 195
QY 218 GGGCGAG 277
DB 196 CGCGAG 255
QY 278 CGCGAG 337
DB 256 CTTGCGAG 315

Db 1417 A 1417

RESULT 8

US-09-105-537-7
Sequence 7, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438051
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 1248
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-7

Query Match 14.9%; Score 74.4; DB 4; Length 1248;
Best Local Similarity 49.5%; Pred. No. 1,1e-05;
Matches 220; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

OY 59 GGGGATCGCTAATGAGACACCCAGTGCATGCGCAACAGAGAGCGCGCGGCGCAT 118
DB 444 GGGGCGCCCTGCGCCCGCCGACGAGCTGCGGAGAGTCCGCGACGACGCGCTGCGGCT 503
OY 119 GCGCGAGAGCGCATGCGGCGCATGCGGCGAGAGGTTGCGCGGCGCATGCGCATG 178
DB 504 GTATGAGCT 563
OY 179 GGTGCG 238
DB 564 GCGGAGAGCG 623
OY 239 GGTGCG 295
DB 624 GCGGAGAGCG 683
OY 296 CCAAGGCGCATGCG 355
DB 684 CTTGCG 743
OY 356 CATTGCG 415
DB 744 GCGGCG 803
OY 416 CAAAGGAGAGCG 475
DB 804 CAAAGGAGAGCG 863
OY 476 GCGGCG 499
DB 864 GCGGCG 887

RESULT 9
US-09-056-556-167/C
Sequence 167, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:

TREATY

ADDRESS: SED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-Apr-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 167:
SEQUENCE CHARACTERISTICS:
LENGTH: 1439 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-056-556-167

Query Match 14.9%; Score 74.4; DB 4; Length 1439;
Best Local Similarity 47.6%; Pred. No. 1,1e-05;
Matches 219; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

OY 8 GCGGATGCTGAGAGCG 67
DB 1413 GCGGATGCTGAGAGCG 1354
OY 68 TAATGAGAGCG 127
DB 1353 TGCGGCG 1294
OY 128 GCGGATGCTGAGAGCG 187
DB 1293 TGCGGCG 1234
OY 188 ACAGGCGCATGCG 247
DB 1233 TGCGGCG 1174
OY 248 CTTGCG 307
DB 1173 TGCGGCG 1114
OY 308 CAAAGGAGAGCG 367
DB 1113 GCGGCG 1054
OY 368 GAGGCG 427
DB 1053 GCGGCG 994
OY 428 GCGGCG 467
DB 993 TGCGGCG 954

RESULT 10
US-09-105-537-3/C
Sequence 3, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.

/ APPLICANT: Liu, H.
 / APPLICANT: Xue, Y.
 / APPLICANT: Zhao, L.
 / TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 / FILE REFERENCE: 600,438US1
 / CURRENT APPLICATION NUMBER: US/09/105,537A
 / CURRENT FILING DATE: 1998-06-26
 / NUMBER OF SEQ ID NOS: 43
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 3
 / LENGTH: 13613
 / TYPE: DNA
 / ORGANISM: Streptomyces venezuelae
 / US-09-105-537-3

Query Match 14.9%; Score 74.4; DB 4; Length 13613;
 Best Local Similarity 49.5%; Pred. No. 1.1e-05;
 Matches 220; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 59 GGGGATGCTAATGAGACACCAACCAAGTGCATCCGACAGAGAGAGAGAGAGAGAGAGAT 118
 DB 12440 GGGGCCCCCTGCGCCGCGCCGACAGCTGCGGAAAGTGGCGACAGACAGAGAGAGAT 12381
 QY 119 GCGGAG 178
 DB 12380 GATCTTGGAT 12321
 QY 179 GGTGCGGCGACAG 238
 DB 12330 GCGGAG 12261
 QY 239 GGTGCAAGGCTTGGCCAG 295
 DB 12260 GCGGAG 12201
 QY 296 CCAAGGAG 355
 DB 12200 CTTCGAG 12141
 QY 356 CATGCTAAG 415
 DB 12140 GCGGAG 12081
 QY 416 CAACGAG 475
 DB 12080 CAACGAG 12021
 QY 476 CGCGCCCGAG 499
 DB 12020 CGACCGCGAG

RESULT 11
 / US-09-177-349-2/c
 / Sequence 2, Application US/09177349
 / Patent No. 6268201
 / GENERAL INFORMATION:
 / APPLICANT: Alland, David
 / APPLICANT: Bloom, Barry R.
 / APPLICANT: Jacobs, Jr., William R.
 / TITLE OF INVENTION: Inhib, Inia AND Inic GENES OF MYCOBACTERIA AND METHODS
 / FILE REFERENCE: 96700/491
 / CURRENT APPLICATION NUMBER: US/09/177,349
 / CURRENT FILING DATE: 1998-10-23
 / NUMBER OF SEQ ID NOS: 14
 / SOFTWARE: Patent Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 5036
 / TYPE: DNA
 / ORGANISM: Mycobacterium tuberculosis
 / US-09-177-349-2

Query Match 14.8%; Score 74; DB 4; Length 5036;
 Best Local Similarity 48.2%; Pred. No. 1.3e-05;
 Matches 209; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 7 ACGGATTCGAG 66
 DB 935 ACGGATTCGAG 876
 QY 67 CAAG 126
 DB 875 CCGGATTCGAG 816
 QY 127 ACGGATTCGAG 186
 DB 815 CCGGATTCGAG 756
 QY 187 CAG 246
 DB 755 CCGGATTCGAG 696
 QY 247 GCTTGGAG 306
 DB 695 CCGGATTCGAG 636
 QY 307 CCAAGGAG 366
 DB 635 CCAAGGAG 576
 QY 367 AGGAG 426
 DB 575 CCGGATTCGAG 516
 QY 427 CCGGATTCGAG 440
 DB 515 AGGAG 502

RESULT 12
 / US-09-320-878-21/c
 / Sequence 21, Application US/09320878A
 / Patent No. 6117659
 / GENERAL INFORMATION:
 / APPLICANT: ASHLEY, Gary
 / APPLICANT: BETLACH, Melanie C.
 / APPLICANT: MCDANIEL, Robert
 / APPLICANT: TANG, Li
 / TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
 / FILE REFERENCE: 30662002120
 / CURRENT APPLICATION NUMBER: US/09/320,878A
 / EARLIER FILING DATE: 1999-05-27
 / EARLIER FILING DATE: 1999-05-27
 / EARLIER FILING DATE: 1998-08-28
 / EARLIER FILING DATE: 1998-08-28
 / EARLIER FILING DATE: 1998-05-06
 / EARLIER FILING DATE: 1997-04-30
 / EARLIER FILING DATE: 1999-02-08
 / EARLIER FILING DATE: 1999-02-08
 / EARLIER FILING DATE: 1998-09-22
 / EARLIER FILING DATE: 1998-09-22
 / EARLIER FILING DATE: 1998-05-28
 / NUMBER OF SEQ ID NOS: 34
 / SOFTWARE: Patent Ver. 2.0
 / SEQ ID NO 21
 / LENGTH: 5970
 / TYPE: DNA
 / ORGANISM: Streptomyces venezuelae
 / US-09-320-878-21

Query Match 14.8%; Score 74; DB 3; Length 5970;
 Best Local Similarity 49.3%; Pred. No. 1.3e-05;

Matches 219; Conservative 1; Mismatches 221; Indels 3; Gaps 1;

[illegible]

RESULT 13 942C-1
 US-08-387-942C-1
 Sequence 1, Application US/08387942C
 Patent No. 5939289
 GENERAL INFORMATION:
 APPLICANT: ERTESVAG, HELGA
 APPLICANT: VALLA, SVEIN
 APPLICANT: SKJAK-BRAEK, GUDMUND
 APPLICANT: LARSEN, BJORN
 TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
 TITLE OF INVENTION: ENCODING MANNITRANAN C-5-EPIMERASE
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
 STREET: P.O. BOX 747
 CITY: CILLS CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22042
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/387,942C
 FILING DATE: 09-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M.
 REGISTRATION NUMBER: 28, 977
 REFERENCE/DOCKET NUMBER: 1809-106P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-205-8000
 TELEFAX: 703-205-8030
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 17588 base pairs

```

1  TYPE: nucleic acid
2  STRANDEDNESS: single
3  TOPOLOGY: linear
4  MOLECULE TYPE: DNA (genomic)
5  ORIGINAL SOURCE:
6  ORGANISM: Azotobacter vinelandii
7  STRAIN: E
8
9  FEATURE:
10 NAME/KEY: CDS
11 LOCATION: 290..1951
12
13 FEATURE:
14 NAME/KEY: CDS
15 LOCATION: 2227..6438
16
17 FEATURE:
18 NAME/KEY: CDS
19 LOCATION: 6702..9695
20
21 FEATURE:
22 NAME/KEY: CDS
23 LOCATION: 9973..12588
24
25 US-08-387-942C-1

```

[illegible]

;; CURRENT APPLICATION NUMBER: US/09/105,537A
;; CURRENT FILING DATE: 1998-06-26
;; NUMBER OF SEQ ID NOS: 43
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 1
;; LENGTH: 15872
;; TYPE: DNA
;; ORGANISM: Streptomyces venezuelae
US-09-105-537-1

Query Match 14.6%; Score 73.2; DB 4; Length 15872;
Best Local Similarity 47.2%; Pred. No. 1.8e-05;
Matches 222; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 8 CGGCATTGCTGAGACGCTGCCAGGCGATGCTAATGAGACGCCGACGAGGCAATCCG 67
DB 4792 CGGCATTGCTGAGACGCTGCCAGGCGATGCTAATGAGACGCCGACGAGGCAATCCG 4851
QY 68 TAATGAGGACACCAACCACTGATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 127
DB 4852 GTACGGGGGCGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 4911
QY 128 CGGCATCCAGGAGATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 187
DB 4912 GAGAGCTGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 4971
QY 188 ACAAGGATGCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 247
DB 4972 CGGCATTGCTGAGACGCTGCCAGGCGATGCTAATGAGACGCCGACGAGGCAATCCG 5031
QY 248 CTTGCGCAACGAGGAG 307
DB 5032 GATCGCGCGCGCGAG 5091
QY 308 CAACGGGAG 367
DB 5092 CGCGCGCGCGCGAG 5151
QY 368 GAGAGCGCTGCGAG 427
DB 5152 GAGAGCGCGCGAG 5211
QY 428 CGCGCAAG 477
DB 5212 CGTCAACGAG 5261

RESULT 15

US-09-025-691-2
; Sequence 2, Application US/09025691
; Patent No. 6069299

GENERAL INFORMATION:

;; APPLICANT: Broadway, Roxanne M.
;; APPLICANT: Harman, Gary E.
;; TITLE OF INVENTION: FUNGUS AND INSECT CONTROL WITH
;; TITLE OF INVENTION: CHITINOLYTIC ENZYMES
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
;; STREET: Clinton Square, P.O. Box 1051
;; CITY: Rochester
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 14603
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/025,691
;; FILING DATE:
;; CLASSIFICATION:

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Goldman, Michael L.
;; REGISTRATION NUMBER: 30,727
;; REFERENCE/DOCKET NUMBER: 19603/20120
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (716) 263-1304
;; TELEFAX: (716) 263-1600
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1294 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
US-09-025-691-2

Query Match 14.5%; Score 72.6; DB 3; Length 1294;
Best Local Similarity 52.0%; Pred. No. 2.3e-05;
Matches 193; Conservative 0; Mismatches 169; Indels 9; Gaps 1;

QY 128 CGGCATCCAGGAGATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 187
DB 317 CACCGGCGCGCGAG 376
QY 188 ACAAGGATGCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 247
DB 377 ACCGGCCACAGGAG 436
QY 248 CTTGCGCAACGAGGAG 307
DB 437 CTTGCGCGAG 496
QY 308 CAACGGGAG 367
DB 497 CAACGGGAG 556
QY 368 GAGAGCGCTGCGAG 427
DB 557 G-----CAATTCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 607
QY 428 CGCGCAAG 487
DB 608 CATCTGCGCGAG 667
QY 488 CATGCGCAACA 498
DB 668 CTTGCGCGACA 678

Search completed: October 10, 2002, 19:46:20
Job time: 1097.73 secs

Page 1

1	153.7	30.7	697	9	AL10383	AL10383E
2	124.6	24.9	179	9	AL041090	AL041090 DKEP434C
3	118.6	23.7	758	10	B18292297	B18292297 6030794282
4	113	22.6	673	10	B1561135	B1561135 603253665
5	98.6	19.7	582	10	BG7243320	BG7243320 602697966
6	94.2	18.8	241	9	A1904151	A1904151 CM-B7043-
7	85.4	17.1	780	10	BG647089	BG647089 102401508
8	80.8	16.2	822	12	CNS04SK	AL305165 Tetraodon
9	79.5	15.9	660	10	BG427000	BG427000 P58561.1
10	79.2	15.8	884	10	BG417341	BG417341 HSNMR002
11	78.4	15.7	990	9	BE2150341	BE2150341 HST0477.1
12	76.6	15.3	682	9	A0176645	A0176645 Tetraodon
13	76.8	15.0	601	12	CNS02EOD	AL183990 Tetraodon
14	73.8	14.8	915	12	CNS01P82	AL2354315 Tetraodon
15	73.4	14.7	401	9	AV438522	AV438522 AV438522
16	73.4	14.7	511	10	BG734444	BG734444 EM1-81_A
17	70.2	14.0	511	10	BG733077	BG733077 6027213808

Fri Oct 11 09:30:46 2002

us-09-489-101a-15_copy_1_500.rst

Page 5

Db	220	CITGAGACTCTGTGACGAGCGTCCGGCTGTGGAMGAAGAAGCGGCCCTTCCTCAAGGCGCGTTTGCCA	279
Oy	76	ACACCACCCATGCATCCTCCCAACGAGAGAACGCCGCCCAAGGCGCATTCGCGAGAGCCGCAATCC	135
Db	280	AGAAAGAGAGATTCTCTCCGACGACTCTGGACACGACACACACACACACATGACACGATATACAGCAGAC	339
Oy	136	AGGCGATCGCCACACGAGAGGTGCCACAGGCGATCTCCCAATGGGTGCCCGCACAGGCA	195
Db	340	ACGAGCCCCGAGAAAGAGGGGCTCCCGCCGCAAGGCGCGCTGTCCAGAAAGGCGCCCGACA	399
Oy	196	TGCGCAATGAGAGCGCCACCAGGCGCATCGCCAATCTGGAGACGCCCTCCACGGCTTGCCA	255
Db	400	GCGACGACGATGACGACACGATGATGACATGTATGACACGACGATGAGCCCCCGCTPAGA	459
Oy	256	ACGGGAGACCGCTTCCTCAGCTTCGGCCCAACGGGAGACGCCGCCCAAGGCGCATTCGCCACAGGAG	315
Db	460	AGGCGAACCCTCGCTGCCGCCGCCGCCGCCGCCGCCCAAGAGGCCCTTCGCCAAGAGAGAG	519
Oy	316	ACGCCACCAAAGGCGANTGGGCAACGAGGTCAACATCCACGCGATCTGCTTAAGAGAGACGCG	375
Db	520	AGTCGCGCGAGATGACGACGATGTACGACGACGACGACGACGACGATGACAGCGAGAGG--	576
Oy	376	TTCAGGGGATCGCTTAACGAGGTGGCGGCCGCCGAGGCGATCGCCAAGAGAGACGCCCGCAGS	435
Db	577	--AGGCGCCCAAGAAAGAGGCGCTGTGCCCTCTCGCGCTGCGCAAGAGGCCCGCGCCAGG	633
Oy	436	GAAATCGCGAGATATGTCCACAGGCGATCGGCAACGAGGACGCCGCCGAGGCGCATCG	492
Db	634	ACGACAGCGATGATGACGAGAGAGACACGATGACACGACGACATGATAGGACATGG	690
RESULT 8			
CNS04SKR/c			
LOCUS	CNS04SKR	822 bp	DNA linear GSS 24-MAY-2000
DEFINITION	Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 007M14 of library H from Tetraodon nigroviridis, genomic survey sequence.		
VERSION	ALJ05165		
KEYWORDS	ALSJ: genome survey sequence.		
SOURCE	GI:8196096		
ORGANISM	Tetraodon nigroviridis. Tetraodon nigroviridis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acantopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.		
REFERENCE	1 (bases 1 to 822) Roest-Criollus,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saulin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis Unpublished 2 (bases 1 to 822) Roest-Criollus,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brotlier,P., Quetier,F., Saulin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence Unpublished 3 (bases 1 to 822)		
JOURNAL	JOURNAL		
AUTHORS	AUTHORS		
TITLE	TITLE		
COMMENT	COMMENT		
FEATURES	FEATURES		
SOURCE	SOURCE		
	Location/Qualifiers		
	1..822 /organism="Tetraodon nigroviridis"		
	/db_xref="taxon:99883"		
	/clone="007M14"		

BASE COUNT	147 a	262 c	203 g	209 t	1 others
ORIGIN					
Query Match	16.2%	Score 80.8	DB 12	Length 822	
Best Local Similarity	50.9%	Pred. No. 0.021			
Matches 217	Conservative	0	Mismatches 206	Indels 3	Gaps 1
QY	20	GGACGCTGCCAGGCGCATATGAGACCGCGACGAGCGATCGTAATGAGGACAC	79		
Db	556	GGAGTAGAGCTAAAGTGTAGCTAACCCGAGTCAAGGTGAAGTGAACCTAACCCGGAGTC	497		
QY	80	CACCAAGTGCATCGCCCAACGAGAACCCGCCCAAGGCATGCG---CGAGACGCGCATCA	136		
Db	496	AGGCTAAGTGCACCTTAACCCGAGTCAAGGTGAAGTGAACCTAACCCGGAGTGAAGCTA	437		
QY	137	GGGATCGCGCAACGAGGAGTTCGCCAGAGGCATCGCCCAATGGGTGCGCCGCAAGGCGCAT	196		
Db	436	AGGTGACCGTAAACCCGAGTCAAGGCTAAAGACGACGCTAACCCCGAGTGAAGGTGA	377		
QY	197	CGCATATGAGGAGACCCACCCAGAGGCATCGCCCAATGGGAGCGCCGTCACGCGCTTCGCGCA	256		
Db	376	CGCTAATCCGGAGTCAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	317		
QY	257	CGGAGGACCGCTCGTCAAGTTCGCCCAAGGGGAGCGCGCCCAAGGATCGGACGCGCAAGGGGA	316		
Db	316	CGTGAAGTCAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	257		
QY	317	CGCCACCAAGGAGCATGGGACGAGGTCACCATCCAGCGATCGTAAAGGAGCGGCGCT	376		
Db	256	GTGAGCTAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	197		
QY	377	CCAGGAGCATCGCTAACGAGTGGCGCGCCAGGCGCATCGCCCAAGGAGGAGCGCGCGGCA	436		
Db	196	CTAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	137		
QY	437	AATGCG	442		
Db	136	TGACGC	131		
RESULT 9					
LOCUS	BE427000				
DEFINITION	BE427000	640 bp	mrna	linear	EST 24-JUL-2000
ACCESSION	PSR6561	ITFC	PSR Wheat Endosperm Library	Triticum aestivum	CNDA
VERSION	BE427000				
KEYWORDS	BE427000.1	GI:9424933			
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					

OY 255 AACGGGAGCCGCTCTCTACGCTTCCGCAACGGGAGCCGCCGCGCATCGCCACAGG 314
 DB 348 AACCTGGAGAGCCCTCTCCGCGCATCCGCGAGGCGCGCGCGCATATCGCCACAGGCG 407
 OY 315 GAGCGCCACCAAGGAGGATGGGCAAGAGATGATCCAGCGCATGCTTAAGAGAGAGCC 374
 DB 408 GAGCGCGCGCGCATATGCTCGAGGCGCTCCGCGCGCATGCTCGCATATGAGCGCGAC 467
 OY 375 GTCCAGGAGCTCGCTAACGAGGTGGCGCGCGCGCATCGCCACAGAGAGCGCGCGAC 434
 DB 468 GTCGCTGCTCTCCGCGCATGAGAGAGAGATGATCTCTATGCAAGAGCATGCGC 527
 OY 435 GGAATGCGCGAGGATGCTGCAAGAGGCA 462
 DB 528 GCAACCTACGACCTCTCATGAGACCA 555

RESULT 11
 BE216934 990 bp mRNA linear EST 03-JUL-2000
 LOCUS BE216934
 DEFINITION EST0477 Triticum aestivum Lambda Zap Triticum aestivum cDNA clone
 JAL_5B_H05_T3 5', mRNA sequence.
 BE216934
 VERSION BE216934.1 GI:8904620
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Triticum.
 1 (bases 1 to 990).

REFERENCE
 AUTHORS Anderson, J.M., Williams, C.E. and Goodwin, S.B.
 TITLE Analysis of an EST database reveals a probable CF2 resistance gene
 JOURNAL hmolog in wheat
 COMMENT Unpublished (2000)
 CONTACT: Anderson, J.M.
 CROP PRODUCTION & Pest Control Research Unit
 USDA-ARS
 1150 Lilly Hall, West Lafayette, IN 47907, USA
 Tel: 765-494-5565
 Fax: 765-496-2926
 Email: janderson@purdue.edu
 Seq primer: T3
 High quality sequence stop: 990.

FEATURES
 source
 1..990
 Location/Qualifiers
 /organism="Triticum aestivum"
 /strain="p29"
 /db_xref="taxon:4565"
 /clone="JAL_5B_H05_T3"
 /clone_lib="Triticum aestivum Lambda Zap"
 /tissue_type="leaf"
 /dev_stage="9 day old seedlings"
 BASE COUNT 182 a 360 c 290 g 158 t

BASE COUNT 182 a 360 c 290 g 158 t
 ORIGIN
 Query Match 15.7%; Score 78.4; DB 9; Length 990;
 Best Local Similarity 50.0%; Pred. No. 0.043;
 Matches 230; Conservative 0; Mismatches 221; Indels 9; Gaps 1;

OY 3 GTTAAAGGCAATGCTGAGAGCGCTGCCAGGCGATCGTATAGAGAGCGGACAGAGG 62
 DB 141 GCTATATCG 200
 OY 63 ATCCGTAATGAGACACACCGCATGCTATCGCCACAGAGAGCGCGCGCGCGCGCG 122
 DB 201 ATCAAGGCG 260
 OY 123 GAGAGCGCGCATCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
 DB 261 GCCAGATCTCTCGAGTCCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 311
 OY 183 GCCGACAGGCG 242

DB 312 CTCACCTGCG 371
 OY 243 CACGCTTGGCG 302
 DB 372 TGGGCTGCG 431
 OY 303 ATGCGCAAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 362
 DB 432 CGCAGCAAGGAGGCG 491
 OY 363 AACGAGAGCG 422
 DB 492 GTCATGCG 551
 OY 423 GACG 462
 DB 552 AAGAGCATGCG 591

RESULT 12
 AUI76645 602 bp mRNA linear EST 21-MAR-2001
 LOCUS AUI76645
 DEFINITION AUI76645 Sugano-Kawakami 5' end enriched cDNA library (OLA) from
 HNI Oryzias latipes cDNA clone OLA14.06c similar to P115282641
 hydroxyproline-rich glycoprotein - maize, mRNA sequence.
 AUI76645
 VERSION AUI76645.1 GI:13424771
 KEYWORDS EST.
 SOURCE Japanese medaka.
 ORGANISM Oryzias latipes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthopterygii; Acanthopterygii; Perciformes; Atherinomorpha;
 Belontiiformes; Atherinomorpha; Oryziatidae; Oryziatidae; Oryzias.
 1 (bases 1 to 602)

REFERENCE
 AUTHORS Naruse, K., Milani, H. and Tanaka, M.
 TITLE Medaka EST Project in University of Tokyo (2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Kiyoshi Naruse
 Department of Biological Sciences
 Graduate School of Science, University of Tokyo
 Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan
 Tel: 81-3-5841-4443
 Fax: 81-3-5841-4410
 Email: naruse@biol.s.u-tokyo.ac.jp
 This clone was isolated from Sugano-Kawakami 5' end enriched cDNA
 library (OLA).

FEATURES
 source
 1..602
 Location/Qualifiers
 /organism="Oryzias latipes"
 /strain="HNI"
 /db_xref="taxon:8090"
 /clone="OLA14.06c"
 /clone_lib="Sugano-Kawakami 5' end enriched cDNA library
 (OLA) from HNI"
 /tissue_type="whole body"
 /dev_stage="adult"

BASE COUNT 132 a 246 c 175 g 49 t
 ORIGIN
 Query Match 15.3%; Score 76.6; DB 9; Length 602;
 Best Local Similarity 49.5%; Pred. No. 0.078;
 Matches 227; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

OY 29 CAGGAGCATGCTAATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 88
 DB 119 CAGAGTGTGCG 178
 OY 89 CATTGCCAAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 145
 DB 179 TGTGGCGAGCG 238

Accession	Definition	LOCUS	Result 13
CNS02EOD	Tetradodon nigroviridis genome survey sequence, 262kH4 of library G from Tetradodon nigroviridis, genomic survey sequence.	681 bp DNA	linear GSS 13-MAY-2000

REFERENCE
AUTHORS
1 (bases 1 to 681)
Roest-Crollius, H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.,

TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i>
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 681)
AUTHORS	Roest-Collins,H., Jallion,O., Dasliya,C., Bouneau,L., Fisher,C.,

TITLE	JOURNAL	REFERENCE
Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence	Unpublished	3 (bases 1 to 681)

JOURNAL COMMENT
Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases this sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigriviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetradon>.

FEATURES	Location/Qualifiers
source	1, .681

```

BASE COUNT      /note="Genoscope sequence ID : C0AG362DD07LP1-end : T7
ORIGIN          74 a      105 c      170 g      282 t
                  50 others

```

Query Match 15.0%; Score 74.8; DB 12; Length 681;

OY	14	TGTTGGAGACCTGCCAGGGCATCGTAAATGAAGAACGCCGACCCAGGGCATTGCTAATA	73
Db	412	TTTTTCGACACACACACACACAAAARACMCTACAAACAAMACMAAOGGCG	35
OY	74	GGAACACCACTGCATGCTGCCAACAGAGAAGCCGCCAGGGCATTCGCGAGACGCT	13
Db	352	CGACAACACAGACGACACGACGACAAAAAGARACGGCGACGACGACGACAA	29
OY	134	CCAGGGCATTCGCCACAGAGAGTGTCACCAGGGCATTCGCAATGGGGTGTCCGACAGGG	19
Db	292	CACGGCAGCACGACGACGACGACGACAAACAAAACRCAACGACGACGACACAA	23
OY	194	CATCGCAATGAGAGCCCAACCCAGGGCATTCGCAACTGGAGGCGCGACAGGCTTGC	25
Db	232	CACACACACACACACAAAAACACACACACRGCAACACACACACACACACACACAA	17
OY	254	CACGGGACGCCCTCTCACTTGCACACAGGGAGGCGCCCAAGGGCATTCGCAACGG	31
Db	172	CAAAAACACACACACACACGCGACACACACACACACACACACACACACACAA	11
OY	314	GGACGCGACCAAGGGCATGGGCAACGAGGTGCACATCCACGGCATTCGTAACAGACGC	37
Db	112	CACACAAAACAA	53
OY	374	CGTCCAGGGCATTCGCTAACGA	394
Db	52	CAASHMACACACACACACAA	32

RESULT 14		
CNS03P82		
LOCUS	915 bp	DNA
CNS03P82		
DEFINITION	Tetradon nigroviridis genome survey sequence pUC-Ori end of clone	GSS 17-MAY-2000

ACCESSION	AL254315
VERSION	AL254315.1
KEYWORDS	GSS; genome survey sequence
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis

REFERENCE
AUTHORS
1 (bases 1 to 915)
Roost-Collins, H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.

TITLE	JOURNAL	REFERENCE
Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i>	unpublished	2 (bases 1 to 915)

TITLE	JOURNAL	REFERENCE
Saurin, W. and Weissenhach, J. Human gene number estimate provided by genome wide analysis using Tetradodon nigroviridis DNA sequence	Unpublished	3 (bases 1 to 915)

FEATURES	JOURNAL COMMENT
Source	Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon . Location/Qualifiers 1..915

```
/organism="Tetraodon nigroviridis"  
/db_xref="taxon:99883"
```


1

1

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

PS Claim 1; SEQ ID NO 8482; 103pp; English.
 VY

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy technique CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (I) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC and responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197/AAS94564 represent novel human CC diagnostic coding sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at http://wipo.int/pub/publ/published_pat_sequences.

SD Sequence 3093 BP; 875 A; 904 C; 862 G; 452 T; 0 other;

Query Match	60.58;	Score 302.4;	DB 23;	Length 3093;
Best Local Similarity	79.88;	Pred. No. 7.2e-38;		
Matches 383; Conservative	0;	Mismatches 91;	Indels 6;	Gaps 2;

QY	24	GTGGCCAGGCGCATGGCTAAATGAGAGAGCGCGACACAGGCGATTCGTATGAGGACACAC	83
Db	1990	GATGCCAAGAAAGCTCCCATGAGAGACCTGTACAGCATCATCTAATACGAGAGCGGCACA	2049
QY	84	CATGTCATTCGCCACAGAGAGAGCGCCGACAGGCGATTCG---CGAGAGCGCATTCACAGGC	140
Db	2050	CACGATATGGCTAACGAGAGATCTGTACAGCATCTGTACAGAGGCGCGCGCCACAAAGGGG	2109
QY	141	ATGCCACAGAGAGAGTGGCCGACAGGCGATTCGCAATGGGGTGGCGCGACAGGCGATTCGC	200
Db	2110	ATCGCCACAGAGAGAGCGCGCGCCACAGCGCATTCGCGACGAGAGAGCGCGCCGACGCGCATTCGC	2169
QY	201	AATGAGAGGACGACCCACAGGCGCATTCGCAACTGGAGAGCGGCTACAGCGCTTCGCGACAGGG	260
Db	2170	AGCGAGAGACCGCGCCACAGGCGCATTCGCGAGAGAGCGCGCCGACGCGCATTCGCGACAGAG	2229
QY	261	GAGCGCGTCTCAGTTTCGCGACAGGAGAGCGCGCCGACAGGCGCATTCGCGACAGGAGAGCGC	320
Db	2230	GAGCGCGCCAGGCGCATTCGCGAGAGAGAGCGCGCCGACAGGCGCATTCGCGACAGGAGAGCGC	2289
QY	321	ACCCAGAGGATGGGCAACGAGAGTTCACCATTCACAGGCGATTCGTTAACGAGAGCGCGTTCAG	380
Db	2290	GGCCAAGGGATTCGCCACAGAGGCGCGCGCCGACAGGCGCATTCGCGAAGGAGAGCGCGCGCCAG	2349
QY	381	GGCATGCGTAACGAGAGTGGCGCCGACAGGCGATTCGCGACAGAGAGAGCGCGCGCCAGGAGATTC	440
Db	2350	GGCATGCGCAAAAGAGGCGCGCGCCGACAGGCGATTCGCGCAACGAGAGAGCGCGCCGACGCGCATTC	2409
QY	441	GC---CGAGATTCGCGACAGGCGATTCGCGACAGAGAGCGCGCGCCGACAGGCGATTCGCGAAC	497
Db	2410	GCCACAGAGAGATCGCGCCACAGGATTCGCGACAGAGAGCGCGCGCCGACAGGCGATTCGCGAATC	2469

RESULT5	
AA572679	
ID	AA572679 standard; cDNA; 2922 BP
XX	
AC	AA572679;

XX 13-FEB-2002 (first entry)
DT
XX
XX
DE DNA encoding novel human diagnostic protein #8483.
DE

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss
 XX
 OS
 PN Homo sapiens.
 PN W0200175067-A2.

PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutation
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 1; SEQ ID No 8483; 103pp; English

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (II) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. The polypeptides and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64157 AAS94364 represent novel human CC diagnostic coding sequences of the invention. CC Note: The sequence data for this patent did not appear in the printed CC specification. Note: This patent was not published in electronic format directly from WIPO CC at http://wipo.int/publ/published_pat_sequences.

Sequence 2922 BP; 833 A; 829 C; 806 G; 454 T; 0 other;

Query Match	53.9%;	Score 269.6;	DB 23;	Length 2922;
Best Local Similarity	75.4%;	Pred. No. 6.3e-33;		
Matches 362;	Conservative 0;	Mismatches 114;	Indels 4;	Gaps 2;

QY 24 GCCTCCCAAGAGCGATCCGTATTGAGAGACGCCGACAGGGGCATCCGTATTAGGACACACCC 83
 Db 1306 GATCCAGAGAAAGCTCCCATGAGACACCCGTACCAACATCACTTAACAGAGATGCTCTCA 1965
 QY 84 CAGTGCATCTGCCACAGAGACGCCGCCACAGGGGCATCCG--CGAGGACGCGATCCAGGCG 140
 Db 1366 CACGATATCTACACAGAGAGCGCTGTCCACGGGCATTCGTAAACAGAGGCGCGCGCCACAGGGC 2025
 QY 141 ATTCGCCACAGAGAGGTTGGCCACAGGGGCATCCGCAATTGGGGTCTCGCGCACAGGGGCATCCGC 200
 Db 2026 ATTCGCCACAGAGAGCGCGCGCCACAGGGGCATTCGCCACAGAGGACGCGCGCCACAGGAATTCGCG 2085

QY 201 AATGAGACCCACCCAGGCGATGCGCAACTGGAGCCCGTCCAGCGCTTCCGCAACGGG 260
 DB 2086 AGCCAGAGACCGCCCGCCAGGCGATGCGCAACGAGGCGCGCCGCGCATCGCAACGAG 2145
 QY 261 GAGCGCTCTCTAGCTTGGCCCAAGGAGCGCCGCCAGGCGATCGCAACGAGGAGCGCC 320
 DB 2146 GATCGCCCGCCAGGCGATGCGCAACGAGGAGCGCGCCGCGCATCGCAACGAGGAGCGCC 2205
 QY 321 ACAGAGGCGATGCGCAACGAGGAGCGCGCATCGCAACGAGGAGCGCGCGCGCGCG 380
 DB 2206 GCGCCAGCGATGCGCAACGAGGAGCGCGCATCGCAACGAGGAGCGCGCGCGCGCG 2265
 QY 361 GCGATCGCTAAGAGGAGGCGCGCCAGGCGATGCGCAACGAGGAGCGCGCGCGCGCG 440
 DB 2266 GCGATCGCTAAGAGGAGGCGCGCCAGGCGATGCGCAACGAGGAGCGCGCGCGCGCG 2325
 QY 441 GCGAGGATGCGCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500
 DB 2326 GCTAATGACACGCTACAGGCA-CGCTAAGAGGAGCGCGCTACAGGAGCGCGCGCGCG 2384

RESULT 6

ABA66861
 ID ABA66861 standard; DNA; 518 BP.

XX ABA66861:

XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #15166.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human foetal liver.

XX Claim 4; SEQ ID NO 15166; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

XX measuring human gene expression in a sample derived from human foetal

XX liver. The single exon nucleic acid probes may be used for predicting,

XX measuring and displaying gene expression in samples derived from human

XX foetal liver. The present sequence is a single exon nucleic acid

XX probe of the invention.

XX Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

Best Local Similarity 75.2%; Pred. No. 5,1e-11;
 Matches 376; Conservative 0; Mismatches 114; Indels 10; Gaps 4;
 QY 5 CTACGCGATTGCTGAGAGACGCTGCGCAGGCGATGCGTAATGAGAGCGCCAGGAGCGAT 64
 DB 23 GATACCTACTTAAGAGATGCTTCAAGATATCTAATACGAGCGCTGTCCAGCGAT 82
 QY 65 GCTAATGAGAGACACCCAGTGCATGCGCAACGAGAGCGCCAGGCGATGCG--- 121
 DB 83 GCTAATGAGAGCGCGCGCCAGGCGATGCGCAACGAGAGCGCGCCAGGCGATGCGCA 142
 QY 122 CGAGAGCG 181
 DB 143 CGAGAGCG 202
 QY 182 GCGCGACAGGCG 241
 DB 203 GCG 262
 QY 242 CCACGCGCTGCGCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
 DB 263 CCATGCGATGCGCAACGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 322
 QY 302 CATGCG 361
 DB 323 CATGCG 382
 QY 362 TAACGAGAGCG 421
 DB 383 TA--ATGACACCGTACAGAGCA-CGCTAAGAGGAGCGCGCTACAGGAGCGCGCGCG 438
 QY 422 GAGCG 478
 DB 439 GAGCG 498
 QY 479 GCG 498
 DB 499 TCTACAGCG 518

RESULT 7

ABA33930
 ID ABA33930 standard; DNA; 518 BP.

XX ABA33930:

XX 23-JAN-2002 (first entry)

DE Probe #12396 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease; ss.

XX Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WP1: 2001-48899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 XX
 PS Claim 4: SEQ ID No 12396; 530pp; English.
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 CC
 XX Sequence 518 BP; 140 A; 170 G; 139 G; 69 T; 0 other;

Query Match 51.5%; Score 257.6; DB 22; Length 518;
 Best Local Similarity 75.2%; Pred. No. 5; Le-31;
 Matches 376; Conservative 0; Mismatches 114; Indels 10; Gaps 4;

OY 5 CTACGGCATTCGTGAGAGCGTCCAGGCGATCGTAAATGAGAGAGCCGACAGGCGAT 64
 DB 23 CAACATCAGTAAGAGAGATGCTCAGAGATATCATTACAGAGAGAGCGTCCAGGCGAT 82
 OY 65 CGCTAATGAG 121
 DB 83 CGCTAATGAG 142
 OY 122 CGAG 181
 DB 143 CGAG 202
 OY 182 CGCGAG 241
 DB 203 CGCGAG 262
 OY 242 CACGAG 301
 DB 263 CAG 322
 OY 302 CAG 361
 DB 323 CAG 382
 OY 362 TAACGAG 421
 DB 383 TAACGAG 438
 OY 422 GAG 478
 DB 439 GAG 498
 OY 479 GAG 498
 DB 499 TCTACAG 518

RESULT 8
 ID AAK15297 standard; DNA: 518 BP.
 AC AAK15297;
 XX AAK15297;
 DT 05-NOV-2001 (first entry)
 XX Human brain expressed single exon probe SEQ ID NO: 15288.

XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 XX Homo sapiens.
 OS
 PN WO200157275-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001MO-US00667.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632486.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WP1: 2001-483446/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4: SEQ ID NO: 15288; 650pp + Sequence listing; English.

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.

Sequence 518 BP; 140 A; 170 G; 139 G; 69 T; 0 other;

Query Match 51.5%; Score 257.6; DB 22; Length 518;
 Best Local Similarity 75.2%; Pred. No. 5; Le-31;
 Matches 376; Conservative 0; Mismatches 114; Indels 10; Gaps 4;

OY 5 CTACGGCATTCGTGAGAGCGTCCAGGCGATCGTAAATGAGAGAGCCGACAGGCGAT 64
 DB 23 CAACATCAGTAAGAGAGATGCTCAGAGATATCATTACAGAGAGAGCGTCCAGGCGAT 82
 OY 65 CGCTAATGAG 121
 DB 83 CGCTAATGAG 142
 OY 122 CGAG 181
 DB 143 CGAG 202
 OY 182 CGCGAG 241
 DB 203 CGCGAG 262
 OY 242 CACGAG 301
 DB 263 CAG 322
 OY 302 CAG 361
 DB 323 CAG 382
 OY 362 TAACGAG 421
 DB 383 TAACGAG 438

CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;

Query Match 47.0%; Score 235; DB 22; Length 495;
 Best Local Similarity 77.3%; Pred. No. 1.3e-27;
 Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

43 ATGAGGAGCGGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 102
 108 ATGAGGAGCGGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 167
 103 AAGCGCGGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 159
 168 ACGGTGTGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 227
 160 CCGAGGATCGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 219
 228 CCGAGGATCGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 287
 220 GCATGCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 279
 288 GCATGCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 347
 280 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 339
 348 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 407
 340 AGGTGACCATGCGGATGCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 399
 408 AGGTGACCATGCGGATGCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 467
 400 CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
 468 CCATATATGACATCGCTAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 494

RESULT 14
 ID AAK02583 standard; DNA; 495 BP.
 AC AAK02583;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 2574.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200157275-A2.
 PN
 PD 09-AUG-2001.
 PD
 PF 30-JAN-2001; 2001WO-US00667.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0613356.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000US-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 PT Penn SG, Hanzel DK, Chen W, Rank DR;

XX
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 2574; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;

Query Match 47.0%; Score 235; DB 22; Length 495;
 Best Local Similarity 77.3%; Pred. No. 1.3e-27;
 Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

43 ATGAGGAGCGGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 102
 108 ATGAGGAGCGGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 167
 103 AAGCGCGGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 159
 168 ACGGTGTGACGAGGATCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 227
 160 CCGAGGATCGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 219
 228 CCGAGGATCGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 287
 220 GCATGCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 279
 288 GCATGCGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 347
 280 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 339
 348 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 407
 340 AGGTGACCATGCGGATGCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 399
 408 AGGTGACCATGCGGATGCGCTAATGAGAGACACCCAGTGTGATGCGCAAGGAGG 467
 400 CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
 468 CCATATATGACATCGCTAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 494

RESULT 15
 ID AAI33949 standard; DNA; 495 BP.
 AC AAI33949;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #2635 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 OS
 PN WO200157272-A2.
 PN
 PD 09-AUG-2001.
 PD
 PF 30-JAN-2001; 2001WO-US00663.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 XX

26-MAY-2000: 2000US-0207456.
PR 30-JUN-2000: 2000US-0608408.
PR 03-AUG-2000: 2000US-0632365.
PR 21-SEP-2000: 2000US-0234687.
PR 27-SEP-2000: 2000US-0236359.
PR 04-OCT-2000: 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX Claim 25: SEQ ID No 2635; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 495 BP; 121 A; 174 C; 129 G; 71 T; 0 other;

Query Match 47.0%; Score 235; DB 22; Length 495;
Best Local Similarity 77.3%; Pred. No. 1,3e-27;
Matches 299; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

QY 43 ATGAGACGCCGACGAGGATCGCTAATGAGACACCCAGTGTGCTGCAAGAGAG 102
DB 108 ATGAGACGCCGACGAGGATCGCTAATGAGACACCCAGTGTGCTGCAAGAGAG 167
QY 103 AAGCCGCCGAGGAGATCGC--CGAGGACGCCATCCAGGAGATCGCCACGAGAGGTTG 159
DB 168 ACGCTGTCCAGGATCGCTAATGAGAGGCCGCCAGGAGATCGCCACGAGAGAGCCG 227
QY 160 CCCAGGATCGGCAATGAGGAGGCCGACAGAGGATCGCAATGAGAGAGCCAGCCAGG 219
DB 228 CCCAGGATCGGCAATGAGGAGGCCGACAGAGGATCGCAATGAGAGAGCCAGCCAGG 287
QY 220 GCATCGCAATGAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAG 279
DB 288 GCATCGCAATGAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAG 347
QY 280 CCAAGGAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCAGC 339
DB 348 CCAAGGAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCAGC 407
QY 340 AGGTCACCATCCAGGATCGCTAATGAGAGAGGCCGCTCCAGGAGGCCGCTAAGAGAGTGG 399
DB 408 AGGATGCCGCTGCAAGGATCGCTAATGAGAGAGGCCGCTCCAGGAGGCCGCTAAGAGATG 467
QY 400 CCGGCCAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCGCTCCAGGAGGCCAGC 426
DB 468 CCATATATGACATCGCTAATGAGAGCCG 494

Search completed: October 10, 2002, 15:17:57
Job time : 94 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 18.7333 Seconds
(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101A-14_COPY_1_500

Perfect score: 500
Sequence: 1 ttagcgcctcgtcgtccccc.....ctcaccgcgcgcgcctcc 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_MN:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	1871	1	US-08-416-478A-1
2	500	100.0	1871	2	US-08-474-988B-1
3	500	100.0	1871	2	US-08-394-442B-1
4	291	58.2	1164	1	US-08-416-478A-5
5	291	58.2	1164	2	US-08-474-988B-5
6	221	44.2	999	1	US-08-394-442B-5
7	221	44.2	999	2	US-08-416-478A-4
8	221	44.2	999	2	US-08-474-988B-4
9	221	44.2	999	2	US-08-394-442B-4
10	78	15.6	718	1	US-08-232-463-14
11	49.2	9.8	53526	3	US-08-658-136-2
12	49.2	9.8	53577	3	US-08-658-136-1
13	47.8	9.6	289	4	US-09-007-005-17
14	47.8	9.6	289	4	US-09-244-796-17
15	42.6	8.5	320	4	US-09-165-264-11
16	41.6	8.3	4488	4	US-08-406-030A-3
17	40.8	8.2	5935	4	US-09-178-973B-17
18	40.8	8.2	5935	4	US-09-354-243B-29
19	40.4	8.1	320	4	US-09-165-264-7
20	39.6	7.9	320	4	US-09-165-264-14
21	39.2	7.8	319	4	US-09-165-264-8
22	39.2	7.8	320	4	US-09-165-264-13
23	39.2	7.8	1954	4	US-09-463-702A-32
24	39.2	7.8	3850	4	US-09-463-702A-1
25	39.2	7.8	31571	1	US-08-323-443B-1
26	39	7.8	1276	4	US-09-177-325-2
27	39	7.8	1276	4	US-09-411-812A-2

c	28	39	7.8	1276	4	US-09-590-113-2
c	29	38.8	7.8	43795	3	US-08-742-185-101
c	30	38.4	7.7	4897	6	5196516-7
c	31	37.8	7.6	318	4	US-09-165-264-12
c	32	37.6	7.5	1456	3	US-09-308-406-1
c	33	37.4	7.5	1721	5	PCT-US96-00994-3
c	34	37.2	7.4	152331	3	US-09-128-155-16
c	35	37	7.4	325	2	US-08-332-766A-11
c	36	37	7.4	12001	1	US-08-458-568A-11
c	37	36.8	7.4	283	5	PCT-US94-04361-19
c	38	36.8	7.4	3408	3	US-09-058-489-14
c	39	36.6	7.3	5322	3	US-09-058-489-13
c	40	36.6	7.3	538	2	US-08-332-766A-24
c	41	36.4	7.3	1773	4	US-08-943-711-215
c	42	36.4	7.3	6246	4	US-08-943-711-640
c	43	36.4	7.3	80246	4	US-09-078-294-4
c	44	36.4	7.3	80595	4	US-09-078-294-3
c	45	36.2	7.2	2580	3	US-09-050-863-2

ALIGNMENTS

RESULT 1
US-08-416-478A-1
Sequence 1, Application US/08416478A
Patent No. 5773578
GENERAL INFORMATION:
APPLICANT: Herkend, Thierry
TITLE OF INVENTION: New Proteins Produced By Human
TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERKEND-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1871 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 231..1724
US-08-416-478A-1
Query Match: 100.0%; Score 500; DB 1; Length 1871;
Best Local Similarity: 100.0%; Pred. No. 6.2e-111;

```

Matches 500: Conservative 0: Mismatches 0: Indels 0: Gaps
QY 1 TCGAGGTGCGGAACTGCCAGCTTTCACAGCTTTCCTGTGATTCGAGCCCTGGATC 60
   1 TCGAGGTGCGGAACTGCCAGCTTTCACAGCTTTCCTGTGATTCGAGCCCTGGATC 60
Db 1 TCGAGGTGCGGAACTGCCAGCTTTCACAGCTTTCCTGTGATTCGAGCCCTGGATC 60
QY 61 CCACCCCAACCCCTCTCCCAAGGCCCTCTCCGTCGTCCTCTCTCTCTCTCTCTCTCT 120
   61 CCACCCCAACCCCTCTCCCAAGGCCCTCTCCGTCGTCCTCTCTCTCTCTCTCTCTCT 120
Db 61 CCACCCCAACCCCTCTCCCAAGGCCCTCTCCGTCGTCCTCTCTCTCTCTCTCTCTCT 120
QY 121 CACCTCCCTCTCTGCAAACTTCTCTTACCCCAACCCCAACCCCAACCCCAACCCCTTTC 180
   121 CACCTCCCTCTCTGCAAACTTCTCTTACCCCAACCCCAACCCCAACCCCAACCCCTTTC 180
Db 121 CACCTCCCTCTCTGCAAACTTCTCTTACCCCAACCCCAACCCCAACCCCAACCCCTTTC 180
QY 181 TTTTTCGACCTCTTTTGGAGGGCTCAGCGCTGCCAAGACATAGAGAGATGTGGAGG 240
   181 TTTTTCGACCTCTTTTGGAGGGCTCAGCGCTGCCAAGACATAGAGAGATGTGGAGG 240
Db 181 TTTTTCGACCTCTTTTGGAGGGCTCAGCGCTGCCAAGACATAGAGAGATGTGGAGG 240
QY 241 CTGAGTCTCTGAGGCTTCTGTCTCTCTCAACCCGCTTTTGGAGGAGCTCCAGTGAAGCTCTCC 300
   241 CTGAGTCTCTGAGGCTTCTGTCTCTCTCAACCCGCTTTTGGAGGAGCTCCAGTGAAGCTCTCC 300
Db 241 CTGAGTCTCTGAGGCTTCTGTCTCTCTCAACCCGCTTTTGGAGGAGCTCTCCAGTGAAGCTCTCC 300
QY 301 AGCAGAGGGCTGAGAGTCCCGGTGTGTGGGGCCAGAGAGGGAGGCTCTGCGCAGCTCCCT 360
   301 AGCAGAGGGCTGAGAGTCCCGGTGTGTGGGGCCAGAGAGGGAGGCTCTGCGCAGCTCCCT 360
Db 301 AGCAGAGGGCTGAGAGTCCCGGTGTGTGGGGCCAGAGAGGGAGGCTCTGCGCAGCTCCCT 360
QY 361 GCAGCCCAACAACTCCCTCCAGAGTCTGAGCTTCTTGGCAAGAGAGGGGTCACTTTGGC 420
   361 GCAGCCCAACAACTCCCTCCAGAGTCTGAGCTTCTTGGCAAGAGAGGGGTCACTTTGGC 420
Db 361 GCAGCCCAACAACTCCCTCCAGAGTCTGAGCTTCTTGGCAAGAGAGGGGTCACTTTGGC 420
QY 421 AGCATCAGCCAGAAAGTGGCCGCGCGCGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 480
   421 AGCATCAGCCAGAAAGTGGCCGCGCGCGCGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 480
Db 421 AGCATCAGCCAGAAAGTGGCCGCGCGCGCGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 480
QY 481 CTCACCCGGAGCGCCCTCC 500
   481 CTCACCCGGAGCGCCCTCC 500
Db 481 CTCACCCGGAGCGCCCTCC 500

RESULT 2
US-08-474-988B-1
: Sequence 1, Application US/08474988B
: Patent No. 5874250
: GENERAL INFORMATION:
: APPLICANT: Herdenc, Thierry
: APPLICANT: Triebel, Frederic
: TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
: TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
: TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROMDY AND NEIMARK, P. L. L. C.
: STREET: 419 Seventh Street, N. W., Suite 400
: CITY: Washington
: STATE: D. C.
: COUNTRY: U. S. A.
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,988B
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/416,478
: FILING DATE: 04-APR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/854,644
: FILING DATE: 08-SEP-1992

```

```

ATTORNEY/AGENT INFORMATION:
1 NAME: YUN, Allen C.
2 REGISTRATION NUMBER: 37,971
3 REFERENCE/DOCKET NUMBER: HERCEND-1B
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (202) 628-5197
6 TELEFAX: (202) 737-3528
7 INFORMATION FOR SEO ID NO. 1:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 1871 base pairs
10 TYPE: nucleic acid
11 STRANDEDNESS: single
12 TOPOLOGY: linear
13 FEATURE:
14 NAME/KEY: CDS
15 LOCATION: 231..1724
16 US-08-474-988B-1
17
18 Query Match 100.0%; Score 500; DB 2; Length 1871;
19 Best Local Similarity 100.0%; Pred. No. 6,2e-111;
20 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
21
22 QY 1 TCAGGCTCCCTGATCTGAGCCAGCCTTCCTCAGCTTTCCCTGATTCGCGCCCTGCTGCATC 60
23 Db 1 TCAGGCTCCGATGATGCGCCAGCTTCCTCAGCTTTCCCTGATTCGCGCCCTGCTGCATC 60
24 QY 61 CCGCCCGCAGCCCTGCTGCTCAAGAGCCCTGCTGCGTCCTCTTCTTGAACCCCTTCCTC 120
25 Db 61 CCGCCCGCAGCCCTGCTGCTCAAGAGCCCTGCTGCGTCCTCTTCTTGAACCCCTTCCTC 120
26 QY 121 CAGCTCCCTGCTGCTGACAAACCTTCCTTCCTTACCCCGCAGCCCGCAGCACTGCGCCCTTTC 180
27 Db 121 CAGCTCCCTGCTGCTGACAAACCTTCCTTCCTTACCCCGCAGCCCGCAGCACTGCGCCCTTTC 180
28 QY 181 TTTTCTGACCTCCCTTTTGGAGGGCTCAGCGCTCCGACACATAGAGAAATGTGGAGG 240
29 Db 181 TTTTCTGACCTCCCTTTTGGAGGGCTCAGCGCTCCGACACATAGAGAAATGTGGAGG 240
30 QY 241 CTCGATTCCTGGGCTTGCTTTTTCGACGCGCCCTTGGGTTGGCTTCAGTGAACCTCTCC 300
31 Db 241 CTCGATTCCTGGGCTTGCTTTTTCGACGCGCCCTTGGGTTGGCTTCAGTGAACCTCTCC 300
32 QY 301 AGCAGAGGGGCTGAAGTCCGAGTGGTGGGCTCAGAGAGGGGGCTCTGACCACCTCCCT 360
33 Db 301 AGCAGAGGGGCTGAAGTCCGAGTGGTGGGCTCAGAGAGGGGGCTCTGACCACCTCCCT 360
34 QY 361 GGAGGCCCAATCCCTCCAGCATCTCAGATCTCCTGAGAGAGAGAGAGGCTACATTGGC 420
35 Db 361 GGAGGCCCAATCCCTCCAGCATCTCAGATCTCCTGAGAGAGAGAGAGGCTACATTGGC 420
36 QY 421 AGCATTCACGACGAGTGGCGCCCGCTGCGCCCGCCGCGCATCCCTGAGCCCGCGGC 480
37 Db 421 AGCATTCACGACGAGTGGCGCCCGCTGCGCCCGCCGCGCATCCCTGAGCCCGCGGC 480
38 QY 481 CTCAGCCGGGGGGGCGCCCTCC 500
39 Db 481 CTCAGCCGGGGGGGCGCCCTCC 500
40
41 RESULT 3
42 US-08-394-442B-1
43 Sequence 1, Application US/08394442B
44 Patent No. 5376877
45 GENERAL INFORMATION:
46 APPLICANT: Hercend, Thierry
47 APPLICANT: Triebel, Frederic
48 TITLE OF INVENTION: New Proteins Produced By Human
49 TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
50 NUMBER OF SEQUENCES: 11
51 CORRESPONDENCE ADDRESSES:
52 ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
53 STREET: 419 Seventh Street, N.W., Suite 400

```

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442B
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HERCEND-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1871 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 231..1724
US-08-394-442B-1

Query Match 100.0%; Score 500; DB 2; Length 1871;
Best Local Similarity 100.0%; Pred. No. 6, 2e-111;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAGGTCGCTGATTCGCCAGCTTCCAGCTTCCCTGATTCGAGCTCGGCTGCTGATC 60
DB 1 TCAGGTCGCTGATTCGCCAGCTTCCAGCTTCCCTGATTCGAGCTCGGCTGCTGATC 60
QY 61 CCTGCCACCCCTCTCTCCAGAGGCTCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTC 120
DB 61 CCTGCCACCCCTCTCTCCAGAGGCTCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTC 120
QY 121 CACCTCCCTCTCTGAGAACTTCTCTTACCCGCCACCCGCCACCCGCCACTGCTTTTC 180
DB 121 CACCTCCCTCTCTGAGAACTTCTCTTACCCGCCACCCGCCACCCGCCACTGCTTTTC 180
QY 181 TTTTCTGACCTCTTTTGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 240
DB 181 TTTTCTGACCTCTTTTGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 240
QY 241 CTCAGTTCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 300
DB 241 CTCAGTTCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 300
QY 301 AGCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
DB 301 AGCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
QY 361 GCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 420
DB 361 GCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 420
QY 421 AGCATGACGACGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 480
DB 421 AGCATGACGACGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 480
QY 481 CTCACCCGCGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 500
DB 481 CTCACCCGCGGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 500

RESULT 4

US-08-416-478A-5
Sequence 5, Application US/08416478A
Patent No. 5773578
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
TITLE OF INVENTION: New Proteins Produced By Human Lymphocytes, DNA Sequences Encoding These Proteins And Their Pharmaceutical And Biological Uses
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 22..1161
US-08-416-478A-5

Query Match 58.2%; Score 291; DB 1; Length 1164;
Best Local Similarity 100.0%; Pred. No. 4, 5e-61;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 210 GCTGCCAGACCTAGAGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 269
DB 1 GCTGCCAGACCTAGAGAGATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 60
QY 270 CCGCTTGGGCTGCTGAGGAGGCTCTCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGG 329
DB 270 CCGCTTGGGCTGCTGAGGAGGCTCTCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGG 329
QY 330 GCCAGAGAGGCTGCTGAGGAGGCTCTCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAG 389
DB 330 GCCAGAGAGGCTGCTGAGGAGGCTCTCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAG 389
QY 390 AGCTTCTGAGAGAGAGGAGGCTCACTTGGAGCATCAGGACGACAGTGGAGGCTGAG 449
DB 390 AGCTTCTGAGAGAGAGGAGGCTCACTTGGAGCATCAGGACGACAGTGGAGGCTGAG 449
QY 450 GCCGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 500
DB 450 GCCGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 500

RESULT 5

US-08-474-988B-5
Sequence 5, Application US/08474988B
Patent No. 5874250
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,988B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/416,478
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: ION, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 22..1161
US-08-474-988B-5

Query Match 58.2% Score 291; DB 2: Length 1164;
Best Local Similarity 100.0%; Pred. No. 4.5e-61;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GCTGCCAGACATAGAGAGATGAGAGGCTCAGTTCTGGGCTTGTGTTCTGCAG 269
DB 1 GCTGCCAGACATAGAGAGATGAGAGGCTCAGTTCTGGGCTTGTGTTCTGCAG 60
QY 270 CCGCTTGGGTGCTCAGTGAAGCTCTCCAGCCAGGGGCTGAGGTCCCGGTGTGTG 329
DB 61 CCGCTTGGGTGCTCAGTGAAGCTCTCCAGCCAGGGGCTGAGGTCCCGGTGTGTG 120
QY 330 GCCCAGAGAGGGGCTCTGCTCCAGGCTCCCGTCCAGCCCAATTCCTCCAGATCTC 389
DB 121 GCCCAGAGAGGGGCTCTGCTCCAGGCTCCCGTCCAGCCCAATTCCTCCAGATCTC 180
QY 390 AGCTTCTGCGAAGAGAGAGGGGTCACTTGGAGCATAGCCAGACAGTGGCCCGCT 449
DB 181 AGCTTCTGCGAAGAGAGGGGTCACTTGGAGCATAGCCAGACAGTGGCCCGCT 240
QY 450 GCCGCCCGGCGCATCCCTGAGCCCGGCGCTACCGCGGCGGCGCTCC 500
DB 241 GCCGCCCGGCGCATCCCTGAGCCCGGCGCTACCGCGGCGGCGCTCC 291

DB 241 GCCGCCCGGCGCATCCCTGAGCCCGGCGCTACCGCGGCGGCGCTCC 291

RESULT 6
US-08-394-442B-5
Sequence 5, Application US/08394442B
Patent No. 5976877
GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
TITLE OF INVENTION: New Proteins Produced By Human
LYMPHOCYTES, DNA Sequences Encoding These Proteins And
TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442B
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HERCEND-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 22..1161
US-08-394-442B-5

Query Match 58.2% Score 291; DB 2: Length 1164;
Best Local Similarity 100.0%; Pred. No. 4.5e-61;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GCTGCCAGACATAGAGAGATGAGAGGCTCAGTTCTGGGCTTGTGTTCTGCAG 269
DB 1 GCTGCCAGACATAGAGAGATGAGAGGCTCAGTTCTGGGCTTGTGTTCTGCAG 60
QY 270 CCGCTTGGGTGCTCAGTGAAGCTCTCCAGCCAGGGGCTGAGGTCCCGGTGTGTG 329
DB 61 CCGCTTGGGTGCTCAGTGAAGCTCTCCAGCCAGGGGCTGAGGTCCCGGTGTGTG 120
QY 330 GCCCAGAGAGGGGCTCTGCTCCAGGCTCCCGTCCAGCCCAATTCCTCCAGATCTC 389
DB 121 GCCCAGAGAGGGGCTCTGCTCCAGGCTCCCGTCCAGCCCAATTCCTCCAGATCTC 180
QY 390 AGCTTCTGCGAAGAGAGAGGGGTCACTTGGAGCATAGCCAGACAGTGGCCCGCT 449
DB 181 AGCTTCTGCGAAGAGAGAGGGGTCACTTGGAGCATAGCCAGACAGTGGCCCGCT 240
QY 450 GCCGCCCGGCGCATCCCTGAGCCCGGCGCTACCGCGGCGGCGCTCC 500
DB 241 GCCGCCCGGCGCATCCCTGAGCCCGGCGCTACCGCGGCGGCGCTCC 291

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: BROMDY AND NEWMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,442B
FILING DATE: 24-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BROMDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HERCEND=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 999 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-394-442B-4

Query Match
Best Local Similarity 44.2%; Score 221; DB 2; Length 999;
Matches 232; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 TCAGGCTGCTGATCTGCCAGCTTCCAGCTTCTCCTCTGATTCGGGCTCTGTCATC 60
DB 768 TCAGGCTGCTGATCTGCCAGCTTCCAGCTTCTCCTCTGATTCGGGCTCTGTCATC 827
QY 61 CCTCCAGCTTCTCTCCAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 828 CCTCCAGCTTCTCTCCAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 867
QY 121 CACCTCCAGCTTCTCTCCAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 888 CACCTCCAGCTTCTCTCCAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 946
QY 181 TTTCTGACCTCTTTGGAGGCTCAGGCTGCCAGACCATAGAGAGATG 233
DB 947 TTTCTGACCTCTTTGGAGGCTCAGGCTGCCAGACCATAGAGAGATG 999

RESULT 10
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)856-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F15
US-08-232-463-14

Query Match
Best Local Similarity 15.6%; Score 78; DB 1; Length 7218;
Matches 12; Conservative 242; Mismatches 132; Indels 0; Gaps 0;

QY 12 GATCTGCCAGCTTCCAGCTTCTCTGATTCGGGCTCTGTCATCCTCCAGGCT 71
DB 1064 GATCTGCCAGCTTCCAGCTTCTCTGATTCGGGCTCTGTCATCCTCCAGGCT 1123
QY 72 TCTCTCCAGGAGGCT 131
DB 1124 TCTCTCCAGGAGGCT 1183
QY 132 CTGACAGAGCT 191
DB 1184 CTGACAGAGCT 1243
QY 192 CTTTGGAGGCT 251
DB 1244 TTTTGGAGGCT 1303
QY 252 GCTTGGCTTCT 311
DB 1304 TTTTGGCTTCT 1363
QY 312 GAGTCCGAGTGTGGAGGCT 371
DB 1364 TTTTGGCTTCT 1423
QY 372 ATCCCTCCAGAGCT 397
DB 1424 TTTTGGCTTCT 1449

RESULT 11
US-08-658-136-2
Sequence 2, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY

```

APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DO$
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GENA-17,8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 53526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-2

Query Match          9.8%; Score 49.2; DB 3; Length 53526;
Best local Similarity 55.2%; Pred. No. 0.0065;
Matches   96; Conservative    0; Mismatches  78; Indels    0; Gaps    0;

QY      50 CTCTGTCATCCCTGCCACCGCTGTCTGCAGAGCCCTCTCCTGGTTCCTTTCTTAGA 109
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 36459 CTCCTCTCTCTCTCTCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 36518

QY      110 ACCCTTCTCTCACCTCCCTCTCTCTGCAGAACTTCTCTTACCACCCCAACCCACCACT 169
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 36519 CCCCCTCTCTCCCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 36578

QY      170 GCCCCTTCTCTTCTCTGACTCTCTTGGAAGGCTGAGCGCTGCGCCAGACCAAT 223
        ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 36579 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 36632


RESULT 12
US-08-658-136-1
Sequence 1, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACIKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
```

```

COMPUTER READABLE FORM:
MEDIM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GENA-17,8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 53577 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

Query Match          9.8%; Score 49.2; DB 3; Length 53577;
Best Local Similarity 55.2%; Pred. No. 0.0065;
Matches 96; Conservative 0; Mismatches 78; Indels 0; Gaps 0.

OY 50 CTTGGCATCCCTCCACCCCTCTCTCCAGAGGCCCTCCCTGCTCCCTCTCTGACG 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36454 CTCGCCCTCTCTCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36513

OY 110 ACCGCTTCTCAGACCTCCCTCTCTCTGACAGACTCTCTCTTACCCCTCCACCCG 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36514 CCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36573

OY 170 GCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36574 CTCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36627

RESULT 13
Sequence 17, Application US/090070058
Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rihc
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
TITLE OF INVENTION: FUSIONS
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
CURRENT FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
EARLIER FILING DATE: 1997-11-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)-(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```


Query Match	9.6%;	Score 47.8;	DB 4;	Length 289;
Best Local Similarity	3.7%;	Pred. No. 0.0039;		
Matches	7;	Conservative	85;	Indels 0;
		Mismatches		Gaps 0;

OY	7	TGCGAGATCTCCACAGCTTTTCACACTTCCGTGGATGCTCGGCCTCTGAGCATCCTCCC	66
Db	247	KGAAAGTYITAYKCYGACGYAAGCYTTCGTGSYNMYSYNMYSYNMYSYNMYSYNMYS	188
OY	67	CACCCCTCTCTCAAGGCCCTCTCCGTGCTTCCTCTTAGAACCCCTTCCTCACCTC	126
Db	187	YNUNSYNMYNSYNMYSYNMYSYNMYSYNMYSYNMYSYNMYSYNMYSYNMYS	128
OY	127	CCTCTCTGACAACCTTCTCTCTTAACCCCCCACCACACACCTGCCCCCTTTCCTTTCT	186
Db	127	YNUNSYNMYNSYNMYSYNMYSYNMYSYNMYSYNMYSYNMYSYNMYSYNMYS	68
OY	187	GACCCCTCTT	197
Db	67	YNUNSYNMYN	57

RESULT 14

```

US-09-244-796-17/C
Sequence 17, Application US/09244796
Patent No. 6281344
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rube
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
TITLE OF INVENTION: FUSIONS
FILE REFERENCE: 00786/350007
CURRENT APPLICATION NUMBER: US/09/244,796
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
EARLIER FILING DATE: 1997-11-06
EARLIER APPLICATION NUMBER: 09/007,005
EARLIER FILING DATE: 1998-01-14
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
NAME/KEY: misc-feature
LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

```

Query Match	9.6%;	Score 47.8;	DB 4;	Length 289;
Best Local Similarity	3.7%;	Pred. No. 0.0039;		
Matches	7;	Conservative	99;	Mismatches 85;
			Indels	0;
			Gaps	0;

[illegible]

RESULT 15
RE=00-16E

```

US-09-165-264-11/C
Sequence 11, Application US/09165264
Patent No. 6197510
GENERAL INFORMATION:
APPLICANT: Vinaayagamorthy, Thuraiayah
TITLE OF INVENTION: Multi-Loci Genomic Analysis
FILE REFERENCE: 44747
CURRENT APPLICATION NUMBER: US/09/165,264
CURRENT FILING DATE: 1996-10-01
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 320
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-11

```

Query Match	8.5%	Score 42.6;	DB 4;	Length 320;
Best Local Similarity	51.3%;	Pred. NO. 0.069;		
Matches 99;	Conservative 0;	Mismatches 94;	Indels 0;	Gaps 0;

[illegible]

Search completed: October 10, 2002, 19:28:21
Job time : 65.7333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 695.733 Seconds
(without alignments)

9699.805 Million cell updates/sec

Title: US-09-489-101A-14_COPY_1-500
Perfect score: 500

Sequence: 1 tcagtgctgctgatctgcc.....ctcaccgcgcgcgcctcc 500

Scoring table: IDENTITY_NUC
Gapop 10.0 Gapext 1.0

Searched: 133736207 seas. 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000000

Post-processing:	Minimum Match 0%
	Maximum Match 100%

Listing first 45 summaries

Database :

```

base :
      EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estin:*
4: em_estmra:*
5: em_estov:*
6: em_estcpl:*
7: em_estro:*
8: em_estcl:*
9:  qb_estcl:*
10: qb_estf:*
11:  qb_hic:*
12:  qb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	298	59.6	586	10	BC898429	BG889429 HOA55-1-2
2	232	46.4	377	10	BC898484	BG908484 HOA19-1-7
3	227	45.4	381	10	BC900314	BG900314 HOA37-1-1
4	150.8	30.2	602	10	BI060106	BI060106 IL3-UT0112
5	149.2	29.8	408	10	BI060032	BI060032 IL3-UT0112
6	69.8	14.0	538	9	AL514267	AL514267 AL514267
7	63	12.6	764	9	AL547921	AL547921 AL547921
8	61.8	12.4	960	12	AG060752	AG060752 Pan trogl
9	60.8	12.2	1222	10	BI16660	BI16660 hsp1601mx
10	59.6	11.9	987	12	CNS00418	AL066537 Drosophi1
11	59	11.8	917	12	AG081214	AG081214 Pan trogl
12	58.6	11.7	796	12	AG10286	AG10286 Pan trogl
13	58.6	11.7	945	12	CNS0076K	AL066860 Drosophi1
14	58.6	11.7	948	12	CNS0016P	AL074439 Drosophi1
15	57.8	11.6	925	12	CNS0031B	AL053031 Drosophi1
16	57.2	11.4	1101	12	CNS0147P	AL018415 Drosophi1
17	56.6	11.3	926	12	AG061568	AG061568 Pan trogl

C	18	56.6	11.3	997	12	CNS0059E	AL060767	Drosophila
C	19	56.2	11.2	932	12	AG134650	AG134650	Pan trogl
C	20	56.2	11.2	949	12	AG078095	AG078095	Pan trogl
C	21	55.8	11.2	966	12	CNS029RG	AL01037	Tetodon
C	22	55.8	11.2	1042	12	B1416533	B1416533	hosp001xe
C	23	55.8	11.2	1101	12	CNS01553	B1404901	hosp001xe
C	24	55.6	11.1	698	10	CB0849405	BG549405	102402500
C	25	55.6	11.1	1183	12	CNS0160R	AL06581	Drosophila
C	26	55.4	11.1	1141	12	BF131888	BF131888	601820838
C	27	55.4	11.1	1236	12	AG126182	AG126182	Pan trogl
C	28	55.2	11.0	703	12	AG1949422	AG194942	Pan trogl
C	29	55.2	11.0	1451	12	BG1291946	BF119496	601884286
C	30	55.5	11.0	724	12	AG171201	AG577101	Pan trogl
C	31	54.8	11.0	955	12	CNS0088E	AL050209	Drosophila
C	32	54.6	10.9	883	13	B1887282	B1887282	ZF637-1.0
C	33	54.6	10.9	1201	12	CNS01550	AL006040	Drosophila
C	34	54.2	10.8	788	12	AG032953	AG032953	Pan trogl
C	35	54.2	10.8	901	12	AG043735	AG043735	Pan trogl
C	36	54.2	10.8	1110	12	BK4759407	BK4759407	ACEN000UR
C	37	54	10.8	663	12	AG122396	AG122396	Pan trogl
C	38	54	10.8	768	9	AG160225	AG160225	Pan trogl
C	39	53.8	10.8	415	9	A1268505	9037c08.x	
C	40	53.8	10.8	672	12	BH140128	BH140128	ZM6BDD000
C	41	53.8	10.8	688	12	CNS039PL	AL083214	Tetodon
C	42	53.8	10.8	776	12	AG081645	AG081645	Pan trogl
C	43	53.8	10.8	995	12	AG135825	AG135825	Pan trogl
C	44	53.8	10.8	952	12	BG435825	BG435825	60247380
C	45	53.8	10.8	1061	12	AG072464	AG072464	Pan trogl

ALIGNMENTS

	LOCUS	586 bp	mRNA	linear	EST 06-NOV-2001
BG98429	DEFINITION				
BG98429	ACCESSION				
BG98429	VERSION				
BG98429.1	KEYWORDS	'mRNA sequence.'			
GI:14308678	SOURCE	HOMO SAPIENS			
human.	ORGANISM	human.			
human sapiens	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	AUTHORS	Kumar,S., Connor,J.R., Dadds,R.A., Halsley,W., Van Horn,M., Mao,J., Salle,G., Mu,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and Lark,M.W.			
Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritis cartilage cDNA libraries	TITLE	Osteoarthritis Cartilage 9 (7), 641-653 (2001)			
21482651	JOURNAL				
MEDLINE	COMMENT	Contact: Sanjay Kumar 0W2109 GlaxoSmithkline 709 Sweedland Road, P.O. Box 1539, King of Prussia, PA 19406, USA Tel: 610-270-7245 Fax: 610-270-5598 Email: sanjay_kumar_1@sk.com Seq primer: 17. Location/Qualifiers 1..586 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="HOA (Human Osteoarthritic Cartilage)" /tissue_type="cartilage" /lab_host="E.coli DH10 B" /note="Vector: pSPORT I; Site-1: SalI; Site-2: NotI; Directional"			
BASE COUNT		80 a	224 c	191 g	90 t others
ORIGIN					

Query Match	59.6%	Score 298	DB 10	Length 586
Best Local Similarity	100.0%	Pred. No. 3.6e-47		
Matches 298	Conservative	0	Mismatches 0	Indels 0
Gaps				
0y	203	GGTCACCGCTGCCAGACACATAGAGAGATGGGAGAGCTCACTTCCTGGCTTCGCTT	262	
Db	83	GGTCACCGCTGCCAGACACATAGAGAGATGGGAGAGCTCACTTCCTGGCTTCGCTT	142	
0y	263	TCGACAGCCGCTTGGTGGTGGCTCCATAGAGCTTCCTACACAGAGGCGTAGAGTCCGGT	322	
Db	143	TCGACAGCCGCTTGGTGGTGGCTCCATAGAGCTTCCTACACAGAGGCGTAGAGTCCGGT	202	
0y	323	GGTGTGGGCGGAGAGAGGGGCTGCTGGCCAGCTCCCTGTGAGCGCCACAAATCCCTCCCA	382	
Db	203	GGTGTGGGCGGAGAGAGGGGCTGCTCCAGCTCCCTGTGAGCGCCACAAATCCCTCCCA	262	
0y	383	GGATTCACACCTTCGGGAGAGACAGAGGTCATCTTGGACATCAACACAGACATGGCC	442	
Db	263	GGATTCACACCTTCGGGAGAGACAGAGGTCATCTTGGACATCAACACAGAGATGGCC	322	
0y	443	GCCCCGTGCGCGCCCGGCGCATCCCTGAGCCCGCGCCCTCACCGCGCGCGCCCTCC	500	
Db	323	GCCCCGTGCGCGCCCGGCGCATCCCTGAGCCCGCGCCCTCACCGCGCGCGCCCTCC	380	
RESULT 2				
EG898484		377 bp	mrna	linear
LOCUS	EG898484			EST 06-NOV-2001
DEFINITION	HOA19-1-A4 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA			
ACCESSION	EG898484			
VERSION	EG898484.1	GI:14308733		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1 (bases 1 to 377)			
	Kumar,S., Connor,J.R., Dadds,R.A., Halsey,W., Van Horn,M., Mao,J., Lark,M.W., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and			
TITLE	Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries			
JOURNAL	Osteoarthrits Cartilage 9 (7), 641-653 (2001)			
MEDLINE	21482651			
COMMENT	Contact: Sanjay Kumar			
	UW2109			
	GlaxoSmithKline			
	709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA			
	Tel: 610-270-7245			
	Fax: 610-270-5598			
	Email: sanjay.kumar-1@sk.com			
	Seq primer: T7			
FEATURES				
source	Location/Qualifiers			
	1..377			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone_lib="HOA (Human Osteoarthritic Cartilage)"			
	/tissue_type="Cartilage"			
	/lab_host="E. coli DH10 B"			
	/note="Vector: pSPORT II; Site_1: SalI; Site_2: NotI; Directional"			
BASE COUNT	53 a 163 c 71 g 89 t 1 others			
ORIGIN				
Query Match	46.4%	Score 232	DB 10	Length 377
Best Local Similarity	99.0%	Pred. No. 1.2e-34		
Matches 232	Conservative	0	Mismatches 1	Indels 0
Gaps				
0y	1	TGAGGTCGCTGATCTGCCAGCTTCCAGCTTCTCTGTGATTCGCGCTGTGATC	60	
Db	1	TGAGGTCGCTGATCTGCCAGCTTCCAGCTTCTCTGTGATTCGCGCTGTGATC	60	

Dd	145	TGAGGGTGCATCTACCTCCCAAGCTTCACAGCTTTCCTCOTGTGATTCCGCGCTTGATCATC	204
Oy	61	CTCTCCGACCCCTCTCTGCCAAGGCGCTCTCTGGTCTCCCTTTCTTAGAAACCCTTCCTC	120
Dd	205	CCCTCCACCCCTCTCTCCAAAGGCGCTCTCTGGTCTCCCTTTCTTAGAAACCCTTCCTC	284
Oy	121	CACCTCCCTCTGCGAAGACTTCCTCTTAACCCCGAACCCCGACACCTGCCCCCTTTC	180
Dd	265	CACCTCCCTCTGCGAAGACTTCCTCTTAACCCCGAACCCCGACACCTGCCCCCTTTC	324
Oy	181	TTTTTCTGACCTCTTTTGAGGGGCTAGCGGCTGCCAGACCATAGAGAGATG	233
Dd	325	TTTTTCTGACCTCTTTTGAGGGGCTAGCGGCTGCCAGACCATAGAGAGATG	377
RESULT 3			
BG900314		381 bp	mRNA linear EST 06-NOV-2003
LOCUS			
DEFINITION			HOA37-1-C11 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA,
ACCESSION			BG900314
VERSION			BG900314.1
KEYWORDS			EST.
SOURCE			human.
ORGANISM			Homo sapiens
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE			1 (bases 1 to 381)
JOURNAL			Kumar, S., Connor, J.R., Dadds, R.A., Halsey, W., Van Horn, M., Mao, J.,
MEDLINE			Sathe, G.V., Mol, P., Agrawal, P., Badger, A.M., Lee, J.C., Cowen, M. and
COMMENT			Lark, M.W. Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries Osteoarthritis Cartilage 9 (7), 641-653 (2001) Contact: Sanjay Kumar UM2109 GlaxoSmithKline 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA Tel.: 610-270-7245 Fax: 610-270-5598 Email: sanjay.kumar-1@sk.com Seq primer: F2 Location/Qualifiers Locus:1.381 ..organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="HOA (human Osteoarthritic Cartilage)" /rname_type="Cartilage" /db_molecule="cDNA B" /note=Vector: pSPORT II; Site_1: SalI; Site_2: NotI; Directional
BASE COUNT			51 a 165 c 70 g 94 t 1 others
ORIGIN			
Query Match			45.4%; Score 227; DB 10; Length 381;
Best Local Similarity			99.6%; Pred. No. 1e-33;
Matches 227:			Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy	1	TCAGGCTGCTGATCTGCCAGGCTTTCACAGCTTCTCTGATTCGGGCTCTGATCATC	60
Dd	154	TCAGGCTGCTGATCTGCCAGGCTTTCACAGCTTCTCTGATTCGGGCTCTGATCATC	213
Oy	61	CTCTCCGACCCCTCTCTGCCAAGGCGCTCTCTGGTCTCCCTTTCTTAGAAACCCTTCCTC	120
Dd	214	CTCTCCGACCCCTCTCTGCCAAGGCGCTCTCTGGTCTCCCTTTCTTAGAAACCCTTCCTC	273
Oy	121	CACCTCCCTCTGCGAAGACTTCCTCTTAACCCCGAACCCCGACACCTGCCCCCTTTC	180
Dd	274	CACCTCCCTCTGCGAAGACTTCCTCTTAACCCCGAACCCCGACACCTGCCCCCTTTC	333
Oy	181	TTTTTCTGACCTCTTTTGAGGGGCTAGCGGCTGCCAGACCATAGAGAGATG	238

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db 334	TTTCTTGACCTCTTTTGGAGGCGCTACGCGTCCGACGACCATANAG 381					
RESULT 4						
LOCUS	Bi060106	602 bp	mRNA	linear	EST 15-JUN-2001	
DEFINITION	IL3-UTRO114-160101-420-A06 UTRO114 Homo sapiens cDNA, mRNA sequence.					
ACCESSION	Bi060106					
VERSION	Bi060106.1	GI:14467633				
KEYWORDS	EST.					
SOURCE	Human					
ORGANISM	Homo sapiens					
REFERENCE	Bukayrova; Melazon; Chorodati; Crtanata; Vertebrata; Euteleostomi; Mammalia; Euteleia; Plimates; Catarhini; Homnidae; Homo.					
AUTHORS	1 (bases 1 to 602) Dias Neto,E., Garcia Correa,R., Verjovsky,I.,Almeida,S., Birones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Colman,G.R., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Goldman,R., SOUZA,V.F.S., Bustos,F., Jongschaap,A., de Souza,V., and M.J. Soares,F., Brentini,R.R., Reis,L.P., de Souza,V., and Simpson,A.J.					
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)					
MEDLINE	20202633 Simpson A J G					
COMMENT	Laboratory of Cancer Genetics Ruda Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel.: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br Project: This database can be seen in the RAPSP/ITCR Human Cancer genome project. This database can be seen in the following URL: (http://www.ludwig.org.br/scripts/gethtml2.pl?i=IL3&t2=IL3-UTRO114-160101-420-A06&t3=2001-01-16&t4=1) Seq primer: puc18 forward High quality sequence stop: 601. Location/Qualifiers 1: 602 bp /db_xref="taxon:9606" /clone_lib="UTRO114" /dev_stage="Adult" /note="Organ: uterus; tumor; Vector: puc18; Site:1: Sma1; Site:2: Sma1; A mini-library was made by cloning products derived from OKSRTS PCR (0.5. Jettens patent application) amplified into the puc18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."					
FEATURES						
source	104 a 228 c 149 g 121 t					
BASE COUNT						
ORIGIN						
Query Match	35.2%	Score 150.8	DB 10:	Length 602:		
Best Local Similarity	95.7%	Score 148.7	DB 10:	Length 602:		
Matches 155:	Conservative 0;	Mismatches 17;	Indels 0;	Gaps 0;		
Db 245	GCAGGGGTCACCTTGAGCAGATCAGCAGCAGAGTGGCCGCC 446					
Db 245	GCAGGGGTCACCTTGAGCAGATCAGCAGCAGAGTGGCCGCC 286					

	RESULT 5	B1060032	408 bp	mRNA	linear	EST 15-JUN-2001
	LOCUS	B1060032	414-160101-419-A10	UT0114	Homo sapiens cDNA, mRNA sequence.	
	DEFINITION	B1060032 Homo sapiens B1060032				
	KEYWORDS	EST.				
	SOURCE	human.				
	ORGANISM	Homo sapiens				
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
	COMMENT	Dias M.A.P., de Garça W. Correa R., Veijolyki-Almeida S., Brites, M.R., Nagel, M.A.F., de Silva W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Cavallieri, P.S., Matsushima, A., Bala, G.S., Simpson, D.H., Brunstein, A., do Oliveira, P.S., Bucher, P., Jonesen, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.U. Shogun sequencing of the human transcriptome with ORF expressed proc Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
	TITLE	Contact: Simpson A.J.U.G.				
	JOURNAL	Laboratory of Cancer Genetics				
	MEDLINE	Ludwig Institute for Cancer Research				
	COMMENT	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-3704-9622 Email: asimpson@ludwig.org.br Fax: +55-11-2707001 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL: http://www.ludwig.org.br/scripts/gethtml2.pl?file=IL3cT2-II3-UT0114 160114-19-A10d13-2001-01-16c14-11 1994_P04 High quality sequence stop: 377.				
	FEATURES	Location/Qualifiers				
	source	1..408				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/clone_lib="UT0114"				
		/dev_stage="Adult"				
		/tissue="Ovarian tumor; Vector: puc18; Site:1: Smat; Site:2: Smat; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application) No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."				
	BASE COUNT	76 a 163 c 90 g 73 t				
	ORIGIN					
	Query Match	29.8%; Score 149.2; DB 10; Length 408;				
	Best Local Similarity	95.1%; Pred. No. 6.4e-19;				
	Matches 154; Conservative	0; Mismatches 8; Indels 0; Gaps 0;				
OY	285	CGACGTGAACTCTTCAGCAGGGAGGTGCCTGCTGGTGCCGCCAAGAGGGCT	344			
DB	125	CTATGTAAGACTCTTCAGCAGGGAGGTGCCTGCTGGTGCCGCCAAGAGGGCT	184			
OY	345	CCTGCCAGCTCCCTGCAGGCCCAACAATCCCCTCCAGAATCATCAGCTTCGACANA	404			
DB	185	CTGCCAGCTCCCTGCAGGCCCAACAATCCCCTCCAGAATCATCAGCTTCGACANA	244			
OY	405	GGAAGGCTACTTGGCAGCATCAGCCAGCAGTAGTGGCCCCGCC	446			
DB	245	GCAAGGCTACTTGGCAGCATCAGCCAGCAGTAGTGGCCCCGCC	286			
RESULT 6	AL514267	538 bp	mRNA	linear	EST 13-FEB-2001	
LOCUS	AL514267	L1_NFL006-pl2	Homo sapiens cDNA clone CL08B007ZD09 3			

V



FILE	TITLE
JOURNAL	Direct Submission
COMMENT	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
FEATURES	Determination of this BAC end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and Est1 libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .
BASE COUNT	165 a 308 c 83 g 185 t 204 others
ORIGIN	Location/Qualifiers
Query Match	11.7%; Score 58.6; DB 12; Length 945;
Best Local Similarity	32.9%; Pred. No. 0.12; Mismatches 139; Indels 0; Gaps 0;
Matches 104; Conservative 74; Mismatches 139; Indels 0; Gaps 0;	
QY 18	CCCAAGCTTCCAGCTTCCTCTGATTCGCGCTCTGTCATCCTCCGACCCCTCTCTC 77
Db 565	CCGSSCCGCGCTCCBCCBCCCCCCCCCKKCCBCTCCBCCCCCTCCGCCCCCCCC 624
QY 78	CAAGCCCTCTCCGCTGCTCCCTTTCTTGAAGACCCCTTCCGACGCTCTCTGAG 137
Db 625	CCCKCCCTTTCBBSBGGCCCTTBSCCCCCBGBCCCCCGCTCKCCCCCGCCCCC 684
QY 138	AAGTCTCTTACCCGCCACCCGACACACATGCCCTTCTCTTTCCTTGAACCTCTTT 197
Db 685	CCKCTCTCKCKKCCCCCGCCGCCCCCCCCCCCGGSGTGKGYTGYBTBGBCB 744
QY 198	GGAGGGCTCAGCGGCGCCAGACCTTAGAGAGAGTGGAGGCTCACTCTCGAGCTTG 257
Db 745	TCTTKTCTTGBGCTBCCSYSOBSKGBBKCKSKTKGTGSKGTSYGKCKTSCCGRTY 804
QY 258	CTGTTTTCGACGCGCTTGGGTGCTCCTCACTGAGCTCTCCACAGCGGCTAGATC 317
Db 805	SYGTCTCKSCGYBKMSCTSTSCCKSYGTGTGGCCTBTBTYTSGCCBCKSKCKSKS 864
QY 318	CCGGTGCTGTGGGCC 333
Db 865	SCKCKCKBSBSSSCS 880
RESULT 14	
CNS001CB	
LOCUS	CNS001CB 948 bp DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence FRT3 end of BAC: BACR36N04 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL074499
VERSION	AL074499.1 GI:4954380
KEYWORDS	GSS.
SOURCE	Fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
COMMENT		determination of this BAC-end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseogawa and Aaron Mammaster BAC library was prepared by Kazutoyo Oseogawa and Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	
FEATURES		location/Qualifiers	
source		1..948	
		/organism="Drosophila melanogaster"	
		/db_xref="taxon:7227"	
		/clone_lib="RPCI-98"	
		/clone="BACR36N04"	
		/note="end : TET3"	
BASE COUNT		130 a 355 c 125 g 146 t 192 others	
ORIGIN			
Query Match		11.7%; Score 58.6; DB 12; Length 948;	
Best Local Similarity		35.4%; Pred. No.0.12;	
Matches 154; Conservative 75; Mismatches 203; Indels 3; Gaps 1;			
QY	69	CCCTCTCTCAAGAGCCCTCTCGTCHCCCTTCTTAGAACCCCTCTCCACTCC	128
		: : : : : : : : : : : : : : : :	
DB	341	YMCCTCYGKCGKCGGGGCTTTTGTVTTCTTTCGGCHTYCTGCGCTYGTCC	400
QY	129	TCTGTGCAACTTCTCTTACCCCAACCCGACACTGCCCCCTTTCCTTCTGA	188
		: : : : : : : : : : : : : : :	
DB	401	GGTTCGGGCTTCCCCCTTGCCTTCCTTTTCCCCCTTCYGGGCTTTTK	460
QY	189	CCCTCTTTTGAAGGCTCAGCGCTCCGACAGACATAGAGAGATGGAGGCTCAGTTC	248
		: : : : : : : : : : : : : : : :	
DB	461	CTTCTCTTKTKKKKKKKKKKKKGTCCGGCTTCTCTGCTCYGGGTCSCSCCTCY	520
QY	249	CTGGGCTTGTCTTCTGCAAGCCGCTTTGGTGAGTCAAGTAAGCTCTCAGCAGG	308
		: : : : : : : : : : : : : : : :	
DB	521	TTSGGGSCCCCTVACVSGSSMTGCCGCTCNCBYGRTKRTTYSGCCSBODGCTGVC	580
QY	309	GCTGAGTTCGCGTGTGTGGGCCAGAGAGGGGC---TCTGCGCCAGCTCCCTGACG	365
		: : : : : : : : : : : : : : : : : :	
DB	581	CCCCBGGCTBTCTGTGKKKGGGSGTGGGCGCCRTYTCGCCCCCTCCGCCCCGCC	640
QY	366	CCCAACATCCCTCCAGAGATCTACGCTTTCGACAGAGAGGGGTCACTTGGCAGAT	425
		: : : : : : : : : : : : : : : :	
DB	641	CCCCGGGCTCCCBGCGGBTBCCYCCCCSCSCCSCSCCBCTGCTTCTKRTGKCCGG	700
QY	426	CAGCCAGACAGTGGCCGCGCTGCGCCCCCGGCAATCCCTGGCCCCGCGCCTTAC	485
		: : : : : : : : : : : : : : : : : :	
DB	701	GCKCYGSGTGTTKGCKKSCCCCCCCCCSCSCCSCSCCSCCCKCKGKSCCTYGS	760
QY	486	CCGGGCGGGCGCCTCC	500
		: : : : : :	
DB	761	CCCTGGCCCGKCYC	775

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 84.9333 Seconds
(Without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101a-14_COPY_1_500

Perfect score: 500

Sequence: 1 tcagagctgcctgctgccc.....ctcacccggcgcgccctcc 500

Scoring table: IDENTITY_NDC

Gapop 10.0, Gapext 1.0

Search: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

1:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	1468	20	AAK06756
2	500	100.0	1629	20	AAK06755
3	500	100.0	1871	12	AAO12869
4	500	100.0	1872	22	AAO11121
5	500	100.0	2279	20	AAK06754
6	291	58.2	1164	12	AAO12871
7	221	44.2	999	12	AAO12870
8	204	40.8	447	19	AAV07462
9	204	40.8	5689	24	ABL33153

C	10	204	40.8	5689	24	ABL34551	Human metastasis a
C	11	106.4	21.3	5689	24	ABL33152	Human immune syste
C	12	106.4	21.3	5689	24	ABL34550	Human metastasis a
C	13	67.2	13.4	3267	23	AAK85793	DNA encoding novel
C	14	67.2	13.4	5912	23	AAK86048	DNA encoding novel
C	15	49.2	10.0	1337	20	AAI17263	Human gene express
C	16	49.2	9.8	53526	19	AAK94101	Human PKD1 gene
C	17	49.2	9.8	53577	17	AAK18551	Human polycystic k
C	18	49.2	9.8	53577	19	AAK94108	Human PKD1 locus b
C	19	48.2	9.6	795	19	AAK55830	FLCA insert stabl
C	20	48	9.6	16033	24	ABL33405	Human immune syste
C	21	47.6	9.5	10732	21	AAK10594	Gene encoding a su
C	22	46	9.2	6338	24	ABL33127	Human immune syste
C	23	45.4	9.1	1503900	22	AAK95240	Human neurigin-1
C	24	45.4	9.1	1503900	22	AAK96733	Human neurigin-1
C	25	44.8	9.0	1050	22	AAK21685	Mutational hot spo
C	26	44.8	9.0	2849	22	AAK21684	Human retinoblast
C	27	44.6	8.9	114955	20	AAK53491	Human adenosine A1
C	28	44.2	8.8	123219	23	AAK88703	Human DNA sequence
C	29	44	8.8	1000	21	AAK02484	Human colon cancer
C	30	43.2	8.6	6167	22	AAK75609	Human immune/haema
C	31	42.6	8.5	320	21	AAK38183	Human immune/haema
C	32	42.6	8.5	12132	22	AAK08699	Primer used in the
C	33	42.4	8.5	989	21	AAK02539	M. carbonacea DNA
C	34	42.4	8.5	32189	22	AAK16341	Human colon cancer
C	35	42.4	8.5	32189	22	AAK17492	Human nervous syst
C	36	42.4	8.5	32189	22	AAK17790	Human nervous syst
C	37	42.4	8.5	32189	22	AAK17790	Human nervous syst
C	38	42.4	8.5	49999	20	AAK23895	Human musculoskele
C	39	42.2	8.4	1218	21	AAK02488	Murine LOBO homolo
C	40	42.2	8.4	6146	22	AAK07552	Human ovarian and
C	41	42.2	8.4	6146	22	AAK07552	Human reproductive
C	42	41.8	8.4	726	22	ABA49384	Human breast liver
C	43	41.8	8.4	726	22	ABA67294	Human breast liver
C	44	41.8	8.4	726	22	ABA34387	Probe #12853 for g
C	45	41.8	8.4	726	22	AAK15731	Human brain expres

ALIGNMENTS

AAK06756	standard; cDNA: 1468 bp.
AAK06756	
26-Apr-1999	(first entry)
Human lymphocyte activation gene 3 splice variant LAG-3V3 cDNA.	
Human lymphocyte activation gene 3; LAG-3V3; splice variant;	
human; immunomodulator; Hashimoto's thyroiditis;	
type 1 diabetes mellitus; multiple sclerosis; Crohn's disease;	
rheumatoid arthritis; allograft rejection;	
graft-versus-host disease; Grave's ophthalmopathy; abortion;	
cerebral malaria; Lyme arthritis; reactive arthritis; hepatitis;	
primary sclerosing colangitis; dermatitis; aplastic anaemia;	
gastroic arthritis; therapy; ss.	
Human sapiens.	
Key	Location/Qualifiers
CDS	297..1313
FT	/*tag= a
W09858059-A1.	
23-DEC-1998.	
03-JUN-1998;	98MO-EP03307.
18-JUN-1997;	97EP-0401404.

XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INRS) INST ROUSSY GUSTAVE.

XX Mastangeli R, Romagnani S, Triebel F;
 DR MPI, 1999-080953/07.
 XX P-PSDB; AAW88361.

PT New lymphocyte activation gene 3 (LAG-3) splice variants - which
 PT can be used as immunomodulators

XX Claim 1; Page 26-27; 49pp; English.

CC This cDNA encodes LAG-3V3 (see AAW88361), a splice variant of human
 CC lymphocyte activation gene 3 (LAG-3) protein. LAG-3V3 derives from
 CC cleavage of the nuclear transcript at a different polyA site
 CC located approximately 170 bp downstream to the 5' end of intron 5.
 CC The retained intron 5 sequence contains an in-frame stop codon. The
 CC resulting mRNA encodes a truncated soluble protein which contains
 CC D1, D2, D3 and 8 new amino acid residues. LAG-3V2 cDNA was
 CC obtained by RT-PCR amplification (see AAX06759 and AAX06762) of
 CC peripheral blood mononuclear cell RNA. The invention provides 3
 CC splice variants (see AAW88359-61) of LAG-3 and nucleotide sequences
 CC encoding them (see AAX06754-56). The LAG-3 variants can be used in,
 CC or for the manufacture of, therapeutic compositions used to treat
 CC immune-related pathologies (claimed), in particular T11-dependent
 CC diseases such as Hashimoto's thyroiditis, type I diabetes mellitus,
 CC multiple sclerosis, Crohn's disease, rheumatoid arthritis, acute
 CC allograft rejection, acute graft-versus-host disease, Grave's
 CC ophthalmopathy, cerebral malaria, Lyme arthritis, reactive arthritis,
 CC HCV-induced chronic hepatitis, primary sclerosing colangitis, contact
 CC dermatitis, unexplained recurrent abortion, aplastic anaemia and
 CC Helicobacter pylori-induced gastric arthritis. LAG-3 variants can
 CC also be used for the manufacture of immunomodulator compounds
 CC (claimed), which can mimic or alter the biological function of
 CC LAG-3.

XX Sequence 1468 BP; 231 A; 550 C; 404 G; 283 T; 0 other;

XX Query Match 100.0%; Score 500; DB 20; Length 1468;

XX Best Local Similarity 100.0%; Pred. NO. 1,9e-105; Indels 0; Gaps 0;

XX Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGGCTGCTGATGCTCCAGCTTTCACAGCTTCTCTGATTCGGCTCTGATC 60
 DB 1 TCAGGCTGCTGATGCTCCAGCTTTCACAGCTTCTCTGATTCGGCTCTGATC 60
 QY 61 CTTCCACCTCTCTGCTGAGGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 120
 DB 61 CTTCCACCTCTCTGCTGAGGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 120
 QY 121 CACCTCTCTCTGCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 DB 121 CACCTCTCTCTGCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
 QY 181 TTTTCTGACCTCTTTTGGAGGCTCAGGCTCCAGACCATAGAGATGTTGGAG 240
 DB 181 TTTTCTGACCTCTTTTGGAGGCTCAGGCTCCAGACCATAGAGATGTTGGAG 240
 QY 241 CTTAGTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 DB 241 CTTAGTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 QY 301 AGCCAGGCTGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 301 AGCCAGGCTGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 361 GCAGCCGCAATCTCCCTTCAGGATCTCAGCTTCTGGAAGAGAGAGGCTCAGTGGC 420
 DB 361 GCAGCCGCAATCTCCCTTCAGGATCTCAGCTTCTGGAAGAGAGAGGCTCAGTGGC 420

OY 421 AGCATCAGCCAGACAGTGGCCCGCCCTGCTCCGCGGCATCCCTGAGCCCGGCC 480
 DB 421 AGCATCAGCCAGACAGTGGCCCGCCCTGCTCCGCGGCATCCCTGAGCCCGGCC 480
 QY 481 CTCACCCGCGCGCCCTCC 500
 DB 481 CTCACCCGCGCGCCCTCC 500

RESULT 2

ID AAX06755 standard; cDNA; 1629 BP.

XX AAX06755;

DT 26-APR-1999 (first entry)

XX Human lymphocyte activation gene 3 splice variant LAG-3V2 cDNA.

DE Lymphocyte activation gene 3; LAG-3; LAG-3V2; splice variant;
 XX human; immunomodulator; Hashimoto's thyroiditis;
 KW type I diabetes mellitus; multiple sclerosis; Crohn's disease;
 KW rheumatoid arthritis; allograft rejection;
 KW graft-versus-host disease; Grave's ophthalmopathy; abortion;
 KW cerebral malaria; Lyme arthritis; reactive arthritis; hepatitis;
 KW primary sclerosing colangitis; dermatitis; aplastic anaemia;
 KW gastric arthritis; therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 297..1565

FT /tag- a

XX W09858059-A1.

XX 23-DEC-1998.

XX 03-JUN-1998; 98MO-EP03307.

XX 18-JUN-1997; 97EP-0401404.

XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INRS) INST ROUSSY GUSTAVE.

XX Mastangeli R, Romagnani S, Triebel F;

XX MPI, 1999-080953/07.

XX P-PSDB; AAW88360.

XX New lymphocyte activation gene 3 (LAG-3) splice variants - which
 XX can be used as immunomodulators

XX Claim 1; Page 24-25; 49pp; English.

CC This cDNA encodes LAG-3V2 (see AAW88360), a splice variant of human
 CC lymphocyte activation gene 3 (LAG-3) protein. LAG-3V2 lacks exon
 CC 6 due to the ligation of the donor site of intron 5 to the acceptor
 CC site of intron 6. There is no frame shift. The LAG-3V2 protein is
 CC a transmembrane protein which does not contain the D4 domain.
 CC LAG-3V2 cDNA was obtained by RT-PCR amplification (see AAX06759 and
 CC AAX06762) of peripheral blood mononuclear cell RNA. The invention
 CC provides 3 splice variants (see AAW88359-61) of LAG-3 and nucleotide
 CC sequences encoding them (see AAX06754-56). The LAG-3 variants can be
 CC used in, or for the manufacture of, compositions used to treat
 CC immune-related pathologies (claimed), in particular T11-dependent
 CC diseases such as Hashimoto's thyroiditis, type I diabetes mellitus,
 CC multiple sclerosis, Crohn's disease, rheumatoid arthritis, acute
 CC allograft rejection, acute graft-versus-host disease, Grave's
 CC ophthalmopathy, cerebral malaria, Lyme arthritis, reactive arthritis,
 CC HCV-induced chronic hepatitis, primary sclerosing colangitis, contact
 CC dermatitis, unexplained recurrent abortion, aplastic anaemia and

DT 24-SEP-2001 (first entry)
XX

Human small cell lung cancer associated gene, LAG-3.

KM Human; small cell lung cancer; therapy; hCAAP; nucleic acid; NA;
 KM melanoma; cancer; colon; breast; head; neck; transitional cancer;
 KM leiomyosarcoma; lymphocyte activation gene-3; LAG-3; synovial sarcoma;
 KM cytosolic; ds.

OS Homo sapiens.
XY

PN W0200153349-A2
 XY

PD 26-JUL-2001

PF 19-JAN-2001; 2001WO-US02015.
xx

PR 21-JAN-2000; 2000US-0489101.

PA (LUDW-) LUDWIG INST CANCER RES.

(COR) CORNELL RES FOUND INC.

PI Stockert E, Scanlan MJ, Jager

DR WPI; 2001-457597/49.

PT Isolated polypeptide, used to treat or prophorose a disorder
PT characterized by expression of a hCAP e-9. cancer, is encoded by an
PX isolated nucleic acid comprising an NA Group 3 or 4 molecule -
XX
Claim 57, Page 108-109, 152pp; English.

CC The invention relates to nucleic acids and encoded polypeptides which
CC are cancer associated antigens expressed in patients afflicted with
CC small cell lung cancer. The molecules provided by the invention can be
CC used in the diagnosis, monitoring, research or treatment of conditions
CC characterised by the expression of one or more cancer associated
CC antigens. The polypeptide is used to treat a disorder characterised by
CC expression of a hCAAP, and determine regression, progression or onset
CC of a condition characterised by expression of an abnormal amount of a
CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
CC are small and non-small cell lung cancer, melanoma, colon, breast, head
CC and neck, or transitional cancer. leiomyosarcoma or synovial sarcoma.
CC The present sequence is a DNA encoding human lymphocyte activation
CC gene-3 (LAG-3). This small cell lung cancer associated gene is
CC designated as NY-SCLC-12.

50 Sequence 1872 BP; 300 A; 673 C; 523 G; 376 T; 0 other;

Query Match	Score	DB	Length
100.0%	500	22	1872

Matches	500;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

QY	1	TGAGGCTGCTGATATGTCGGCCAGCTTTCACGCTTTCGTCGAGATTCGCGCTCTGGTATG	60
Db	1	TGAGGCTGCTGATATGTCGGCCAGCTTTCACGCTTTCGTCGAGATTCGCGCTCTGGTATG	60
QY	61	CGCCCCACACCCCTCTCCCAAGGCGCTCTCGTGCATCCCTCTGTATGAACCCCTCTGC	120
Db	61	CGCCCCACACCCCTCTCCCAAGGCGCTCTCGTGCATCCCTCTGTATGAACCCCTCTGC	120
QY	121	GACGTCGCTCTGTGAGACACTTCTCTTTCACCCCAACCCCAACACTGCGCCCTTTC	180
Db	121	GACGTCGCTCTGTGAGACACTTCTCTTTCACCCCAACCCCAACACTGCGCCCTTTC	180
QY	181	TTTTTCTGACCTCTTTTGGAGGCGCTCGAGCGCTGCCAGACCATAGAGAGATGTGGAGG	240
Db	181	TTTTTCTGACCTCTTTTGGAGGCGCTCGAGCGCTGCCAGACCATAGAGAGATGTGGAGG	240
QY	241	CTCAGTTCCTGGGCTTCTGTTTCTGAGCGCGCTTTTGGGTGGCTCTCAGTGAAGCTCTCC	300
Db	241	CTCAGTTCCTGGGCTTCTGTTTCTGAGCGCGCTTTTGGGTGGCTCTCAGTGAAGCTCTCC	300

[illegible]

RESULT 5

ID AAX06754 standard; cDNA; 2279 BP.

AC AAX06754;

DT 26-APR-1999 (first entry)

DE Human lymphocyte activation gene 3 splice variant LAG-3V1 cDNA.

KW Lymphocyte activation gene 3; LAG-3; LAG-3V1; splice variant;

KW type I diabetes mellitus; multiple sclerosis; Crohn's disease;

KW graft-versus-host disease; Grave's ophthalmopathy; abortion;

KW primary sclerosing colangitis; dermatitis; aplastic anaemia;

XX

XX

FT	CDS
----	-----

XX

XX

XX 7
C
C
D
C
t
v
c

[illegible]

TO DON TSC/ 3/EF 040104,
XX
XX

PA (INST) ARS APPLIED RES SISTEMS HOLDING NV.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE

XX (INSE) INDI NOUSSI GOSIHE.

PI Maslach, R., Kromann, S., Liebel, R.
XX

DR WPL; 1999-080953/
DB P-PSDB; AA88359

XX New lymphocyte activation zone 3 (LAC-3) enzyme variants - which

PS Claim 1; Page 21-22; 49pp; English.

CC This cDNA encodes LAG-3V1, an splice variant of human
CC lymphocyte activation gene 3 (AA068359). LAG-3V1 is derived
CC from the retention of intron 4, i.e. cleavage at the donor and
CC acceptor sites flanking intron does not occur. An in-frame stop
CC codon located after 8 codons in the retained intron 4 leads to a 36
CC kDa truncated soluble LAG-3V1 protein containing D1, D2 and 8 new
CC amino acid residues. LAG-3V1 cDNA was obtained by RT-PCR
CC amplification (see AA06757-58) of peripheral blood mononuclear cells.

CC RNA. The invention provides 3 splice variants (see AAM8359-61) of
 CC LAG-3 and nucleotide sequences encoding them (see AAX06754-56). The
 CC LAG-3 variants can be used in, or for the manufacture of,
 CC pharmaceutical or therapeutic compositions which are used to treat
 CC immune-related pathologies (claimed), in particular Th1-dependent
 CC diseases such as Hashimoto's thyroiditis, type 1 diabetes mellitus,
 CC multiple sclerosis, Crohn's disease, rheumatoid arthritis, acute
 CC allograft rejection, acute graft-versus-host disease, Grave's
 CC ophthalmopathy, cerebral malaria, Lyme arthritis, reactive arthritis,
 CC HCV-induced chronic hepatitis, primary sclerosing colangitis, contact
 CC dermatitis, unexplained recurrent abortion, aplastic anaemia and
 CC Helicobacter pylori-induced gastric antritis. LAG-3 variants can
 CC also be used for the manufacture of immunomodulator compounds
 CC (claimed), which can mimic or alter the biological function of
 CC LAG-3.

XX Sequence 2279 BP; 360 A; 832 C; 589 G; 498 T; 0 other;

Query Match 100.0%; Score 500; DB 20; Length 2279;

Best Local Similarity 100.0%; Pred. No. 2.1e-105; Mismatches 0; Indels 0; Gaps 0;

Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 TCAGGCTCCGATCCGACGCTTTCAGTCTTCCTGATCCGCGCTGCTGATC 60
 1 TCAGGCTCCGATCCGACGCTTTCAGTCTTCCTGATCCGCGCTGCTGATC 60
 61 CTTCCCAACCTCTCTCCAGAGCCCTCTGCTGCTCTCTCTCTCTCTCTCTC 120
 61 CTTCCCAACCTCTCTCCAGAGCCCTCTGCTGCTCTCTCTCTCTCTCTCTC 120
 DB 121 CACCTCCCTCTCTGAGAACTTCTCTTACCCGACCCGACGCTGCTGCTTTC 180
 121 CACCTCCCTCTCTGAGAACTTCTCTTACCCGACCCGACGCTGCTGCTTTC 180
 DB 181 TTTTCTGACCTCTTGTGAGAGGCTCAGCGCTGCCAGACATAGAGAGATGAGG 240
 181 TTTTCTGACCTCTTGTGAGAGGCTCAGCGCTGCCAGACATAGAGAGATGAGG 240
 DB 241 CTTCACTTCTGAGGCTGCTGTTTCTGAGAGCCGCTTGTGAGGCTGCTGAG 300
 241 CTTCACTTCTGAGGCTGCTGTTTCTGAGAGCCGCTTGTGAGGCTGCTGAG 300
 DB 301 AGCCAGGGGCTGAGGCTCCGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 360
 301 AGCCAGGGGCTGAGGCTCCGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 360
 DB 361 GCAGCCCAACATCCCTCCAGATCTCAGCTTCTGCGAAGAGAGAGGCTCATT 420
 361 GCAGCCCAACATCCCTCCAGATCTCAGCTTCTGCGAAGAGAGAGGCTCATT 420
 DB 421 AGCATCAGCAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 421 AGCATCAGCAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 481 AGCATCAGCAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 481 AGCATCAGCAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 481 CTTACCCGCGGCGGCTTCC 500
 481 CTTACCCGCGGCGGCTTCC 500

RESULT 6

AAQ12871 standard; DNA; 1164 BP.

AAQ12871;

15-OCT-1991 (first entry)

Polyhedrin-soluble LAG-3 fusion construct.

human cytotoxic T-lymphocyte; LAG-3S; HIV; CD4; AIDS; SS.

XX Homo sapiens.

Key Location/Qualifiers
 sig_peptide 22..105
 mat_peptide 106..1161
 /tag= a
 /tag= b
 /product= LAG-3S fused to polyhedrin fragment

MO9110682-A.

25-JUL-1991.

08-JAN-1991; 91WO-FR00009.

08-JAN-1990; 90FR-0000126.

(INRA) INSERM INST NAT SANTE.

(INSR) INST ROUSSY G.

Herend T, Triebel F;

WPI: 1991-237986/32.

P-PSDB; AAR13272.

Human lymphocyte membrane protein and DNA encoding it (LAG-3) -
 useful for antibodies and pharmaceutical products used to treat
 HIV-infected humans

Example 11; Page 49-50; 69pp; French.

FDC cDNA was isolated from a natural cytotoxic lymphocyte cDNA
 library constructed in lambda gt10. It encodes LAG-3, a
 membrane-bound protein. A soluble form (LAG-3S) of the protein can
 be produced in insect cells (i.e. Spodoptera frugiperda Sf9 cells)
 by removing the transmembrane region-coding sequence. This
 construct encodes the first 3 extracellular Ig-like domains of
 LAG-3 (V domain and two C2 domains) fused to a 17 amino acid
 C-fragment of the insect protein polyhedrin.
 See also MA012869-70.

Sequence 1164 BP; 182 A; 426 C; 340 G; 216 T; 0 other;

Query Match 58.2%; Score 291; DB 12; Length 1164;

Best Local Similarity 100.0%; Pred. No. 1.4e-57; Mismatches 0; Indels 0; Gaps 0;

Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 210 GCTGCCAGACCATAGAGAGATGTGGAGGCTCAGTCTCTGCTGCTGCTGCTG 269
 1 GCTGCCAGACCATAGAGAGATGTGGAGGCTCAGTCTCTGCTGCTGCTGCTGCTG 269
 DB 270 CGCTTTGGGTGCTCAGTGAAGGCTTCCAGCCAGGGGCTGAGGCTCCGCTGCTG 329
 270 CGCTTTGGGTGCTCAGTGAAGGCTTCCAGCCAGGGGCTGAGGCTCCGCTGCTG 329
 DB 61 CGCTTTGGGTGCTCAGTGAAGGCTTCCAGCCAGGGGCTGAGGCTCCGCTGCTG 120
 61 CGCTTTGGGTGCTCAGTGAAGGCTTCCAGCCAGGGGCTGAGGCTCCGCTGCTG 120
 DB 330 GCCCAGAGGGGCTCTCCGACGCTCCCTGACGCTCCCATCTCCCTCCAGATCTC 389
 330 GCCCAGAGGGGCTCTCTCCGACGCTCCCTGACGCTCCCATCTCCCTCCAGATCTC 389
 DB 121 GCCCAGAGGGGCTCTCTCCGACGCTCCCTGACGCTCCCATCTCCCTCCAGATCTC 180
 121 GCCCAGAGGGGCTCTCTCCGACGCTCCCTGACGCTCCCATCTCCCTCCAGATCTC 180
 DB 390 AGCTTCTGCGAAGAGAGAGGCTTGTGAGCATCAAGCCAGAGAGAGGCTCCGCT 449
 390 AGCTTCTGCGAAGAGAGAGGCTTGTGAGCATCAAGCCAGAGAGAGGCTCCGCT 449
 DB 181 AGCTTCTGCGAAGAGAGAGGCTTGTGAGCATCAAGCCAGAGAGAGGCTCCGCT 240
 181 AGCTTCTGCGAAGAGAGAGGCTTGTGAGCATCAAGCCAGAGAGAGGCTCCGCT 240
 DB 450 GCCGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
 450 GCCGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
 DB 241 GCCGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 291
 241 GCCGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 291

RESULT 7

AAQ12870 standard; DNA; 999 BP.

AAQ12870;

15-OCT-1991 (first entry)

KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 OS Homo sapiens.
 XX MO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI: 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 1126; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 SQ Sequence 5689 BP; 1238 A; 120 C; 1590 G; 2741 T; 0 other;
 XX
 XX
 Query Match 40.8%; Score 204; DB 24; Length 5689;
 Best Local Similarity 81.2%; Pred. No. 1.5e-37;
 Matches 237; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 OY 1 TCAGGCTGCTGATCTGCCAGCTTTCAGCTTCTCTGATTCGGGCTTGGTCATC 60
 DB 570 TCAAACTACCTATCTATCCCAACTTTCCTCTAAATTCGACCTTAATCATC 511
 OY 61 CCTCCACACCTCTCTCAGAGCCCTCTCTGTCCTCTCTCTAGAACCTTCTC 120
 DB 510 CCTCCACACCTCTCTCAGAGCCCTCTCTGTCCTCTCTCTAGAACCTTCTC 451
 OY 121 CACCTCCCTCTCTGAGAACTTCTCTTACCCGCCACCCACACCTGCCCCCTTC 180
 DB 450 CACCTCCCTCTCTGAGAACTTCTCTTACCCGCCACCCACACCTGCCCCCTTC 391
 OY 181 TTTTCTGACCTCTCTTTGGAGGCTCAGCGCTGCCAGACCATAGAGAGATGGGAGG 240
 DB 390 TTTTCTGACCTCTCTTTTAAAGCTCAAGCTACCAAGCAATAAATAAATAA 331
 OY 241 CTGAGTCTCTGGGCTCTGCTTTCTGACGCGCTTTGGTGGCTCTCAAGTAA 292
 DB 330 CTCAATTCCTAACTTACTATTCTACACCGCTTAAATTAATCACTCAATAA 279

RESULT 10
 ABL34551/c
 ID ABL34551 standard; DNA; 5689 BP.
 XX
 AC ABL34551;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human metastasis associated gene SEQ ID NO: 104.
 XX

KW Metastasis associated gene; cytostatic; gene therapy; cancer;
 KW cytosine methylation; gene; ds.
 OS Homo sapiens.
 XX MO200177376-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP03970.
 XX
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI: 2002-010922/01.
 XX
 PT New nucleic acid derived from chemically treated metastasis genes,
 PT useful for diagnosis of cancers by analysis of cytosine methylation,
 PT also for treatment -
 XX
 PS Claim 1; SEQ ID NO 104; 23pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of human metastasis associated
 CC genes which are modified by cytosine methylation. The sequences can be
 CC used in the diagnosis and treatment of cancer. The present sequence is
 CC one of the genes of the invention.
 CC
 SQ Sequence 5689 BP; 1238 A; 120 C; 1590 G; 2741 T; 0 other;
 XX
 XX
 Query Match 40.8%; Score 204; DB 24; Length 5689;
 Best Local Similarity 81.2%; Pred. No. 1.5e-37;
 Matches 237; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 OY 1 TCAGGCTGCTGATCTGCCAGCTTTCAGCTTCTCTGATTCGGGCTTGGTCATC 60
 DB 570 TCAAACTACCTATCTATCCCAACTTTCCTCTAAATTCGACCTTAATCATC 511
 OY 61 CCTCCACACCTCTCTCAGAGCCCTCTCTGTCCTCTCTCTAGAACCTTCTC 120
 DB 510 CCTCCACACCTCTCTCAGAGCCCTCTCTGTCCTCTCTCTAGAACCTTCTC 451
 OY 121 CACCTCCCTCTCTGAGAACTTCTCTTACCCGCCACCCACACCTGCCCCCTTC 180
 DB 450 CACCTCCCTCTCTGAGAACTTCTCTTACCCGCCACCCACACCTGCCCCCTTC 391
 OY 181 TTTTCTGACCTCTCTTTGGAGGCTCAGCGCTGCCAGACCATAGAGAGATGGGAGG 240
 DB 390 TTTTCTGACCTCTCTTTTAAAGCTCAAGCTACCAAGCAATAAATAAATAA 331
 OY 241 CTGAGTCTCTGGGCTCTGCTTTCTGACGCGCTTTGGTGGCTCTCAAGTAA 292
 DB 330 CTCAATTCCTAACTTACTATTCTACACCGCTTAAATTAATCACTCAATAA 279

RESULT 11
 ABL33152
 ID ABL33152 standard; DNA; 5689 BP.
 XX
 AC ABL33152;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 1125.
 XX
 DE Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW

KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

PF 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation

PS Claim 1; SEQ ID NO 1125; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

CC Sequence 5689 BP; 1151 A; 120 C; 1710 G; 2708 T; 0 other;

Query Match 21.3%; Score 106.4; DB 24; Length 5689;
 Best Local Similarity 60.3%; Pred. No. 3.5e-15;
 Matches 176; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 1 TCAGGCTGCGATGTCGCCAGCTTCCGATTCGCGGCTGCGTATC 60

DB 5120 TTAGGTTGTTGATTTGTTAGTTTGTGTTTGTGTTTGTGTTTGTGTTAT 5179

QY 61 CTTCCCAACCTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 120

DB 5180 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5239

QY 121 CACCTCCCTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 180

DB 5240 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5299

QY 181 TTTTCTGACCTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 240

DB 5300 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5359

QY 241 CTCAGTTCTGCGGCTTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 292

DB 5360 TTTAGTTTGTGGTTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTAA 5411

RESULT 12

ID ABL34550 standard; DNA; 5689 BP.

AC ABL34550;

DT 26-MAR-2002 (first entry)

DE Human metastasis associated gene SEQ ID NO: 103.

KW Metastasis associated gene; cytostatic; gene therapy; cancer;
 KW cytosine methylation; gene; ds.

OS Homo sapiens.

PN WO200177376-A2.

PD 18-OCT-2001.

PF 06-APR-2001; 2001WO-EP03970.

PR 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-010922/01.

PT New nucleic acid derived from chemically treated metastasis genes,
 PT useful for diagnosis of cancers by analysis of cytosine methylation,
 PT also for treatment

PS Claim 1; SEQ ID NO 103; 23pp + Sequence Listing; English.

CC The present invention provides a number of human metastasis associated
 CC genes which are modified by cytosine methylation. The sequences can be
 CC used in the diagnosis and treatment of cancer. The present sequence is
 CC one of the genes of the invention.

CC Sequence 5689 BP; 1151 A; 120 C; 1710 G; 2708 T; 0 other;

Query Match 21.3%; Score 106.4; DB 24; Length 5689;
 Best Local Similarity 60.3%; Pred. No. 3.5e-15;
 Matches 176; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 1 TCAGGCTGCGATGTCGCCAGCTTCCGATTCGCGGCTGCGTATC 60

DB 5120 TTAGGTTGTTGATTTGTTAGTTTGTGTTTGTGTTTGTGTTTGTGTTAT 5179

QY 61 CTTCCCAACCTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 120

DB 5180 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5239

QY 121 CACCTCCCTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 180

DB 5240 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5299

QY 181 TTTTCTGACCTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 240

DB 5300 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5359

QY 241 CTCAGTTCTGCGGCTTCTCTCCAGGCTTCCGATTCGCGGCTGCGTATC 292

DB 5360 TTTAGTTTGTGGTTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTAA 5411

RESULT 13

ID AAS85793/C standard; CDNA; 3267 BP.

AC AAS85793;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #21597.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

```
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG21606.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 1; SEQ ID No 21597; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3267 BP; 572 A; 914 C; 970 G; 811 T; 0 other;
XX
XX Query Match 13.4%; Score 67.2; DB 23; Length 3267;
XX Best Local Similarity 95.8%; Pred. No. 3e-06; 3; Indels 0; Gaps 0;
XX Matches 69; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 429 CCAACAGTGGCGCGCGCGCGCGCGCGCATCCCTGCGCCCGGCGCCCTACCCG 488
XX |||
XX Db 2250 CGCGCCAGTGGCGCGCGCGCGCGCGCGCATCCCTGCGCCCGGCGCCCTACCCG 2191
XX
XX 489 GCGGCGCCCTCC 500
XX |||
XX Db 2190 GCGGCGCCCTCC 2179
XX
XX RESULT 14
XX AAS86048
XX ID AAS86048 standard; cDNA; 5912 BP.
XX AC AAS86048;
XX AC AAS86048;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #21852.
```

```
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG21861.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 1; SEQ ID No 21852; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 5912 BP; 1540 A; 1587 C; 1551 G; 1234 T; 0 other;
XX
XX Query Match 13.4%; Score 67.2; DB 23; Length 5912;
XX Best Local Similarity 95.8%; Pred. No. 3.3e-06; 3; Indels 0; Gaps 0;
XX Matches 69; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 429 CCAGACAGTGGCGCGCGCGCGCGCGCGCGCATCCCTGCGCCCGGCGCCCTACCCG 488
XX |||
XX Db 1018 CGCGCCAGTGGCGCGCGCGCGCGCGCGCATCCCTGCGCCCGGCGCCCTACCCG 1077
XX
XX 489 GCGGCGCCCTCC 500
XX |||
XX Db 1078 GCGGCGCCCTCC 1089
XX
XX RESULT 15
XX AA217263
XX ID AA217263 standard; cDNA; 1337 BP.
XX AC AA217263;
XX AC AA217263;
XX DT 12-OCT-1999 (first entry)
```

XX Human gene expression product cDNA sequence SEQ ID NO:4735.

XX Human: gene: gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

OS Homo sapiens.

XX MO9938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99MO-US01619.

XX 03-APR-1998; 98US-0080666.

XX 28-JAN-1998; 98US-0072910.

XX 24-FEB-1998; 98US-0075954.

XX 31-MAR-1998; 98US-0080114.

XX 03-APR-1998; 98US-0080515.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;

XX Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;

XX Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;

XX Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

XX Stache-Crain B, Sudduth-Klinger J, Williams LT;

XX WPI; 1999-494092/41.

XX Novel human genes and their expression products which are

XX differentially expressed in different cell types

XX Claim 1: Page 2250-2251; 2479pp; English.

XX The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AA12532 to AA17779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AA12532 to AA17779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purpose, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensic, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.

XX Sequence 1337 BP; 42 A; 577 C; 27 G; 22 T; 669 other;

XX Query Match 10.0%; Score 49.8; DB 20; Length 1337;

XX Best Local Similarity 26.5%; Pred. No. 0.025;

XX Matches 132; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

XX 2 CAGGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 61

XX 592 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651

XX 62 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 121

XX 652 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 711

QY 122 ACCTCCCTCTGAGAACTTCTCTTACCCTCCACACCCACACCTGCTTCTCT 181
 DB 712 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 QY 182 TTTCTGACCTCTTTTGGAGGCTCAGCGCTGCCAGACCATAGAGATGTTGGAGGC 241
 DB 772 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 QY 242 TCAGTTCCTGAGCTTGTCTGTTCTGACGCGCTTTGGTGGTGCATGAGCTTCTCA 301
 DB 832 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 QY 302 GCGAGGGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
 DB 892 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 QY 362 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 DB 952 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 QY 422 GCATCAGCAGACAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 481
 DB 1012 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651
 QY 482 TCACCG 500
 DB 1072 CACGCTGCTGATGTCGACGTTTCCAGCTTTCCTGATTCGCGCTTGGTCATCC 651

Search completed: October 10, 2002, 15:17:48
 Job time: 102 secs

Fri Oct 11 09:30:41 2002

us-09-489-101a-13_copy_1_500.rni

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 18.7333 Seconds
(without alignments)
6536.054 Million cell updates/sec

Title: US-09-489-101a-13_COPY_1_500
Perfect score: 500
Sequence: 1 gcccgaccggaagtgag.....gcgacttgcagagactcc 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/prodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCFUS.COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	60	12.0	2580 3	US-09-050-863-2 Sequence 2, Appl
2	60	12.0	2580 4	US-09-359-081-2 Sequence 1, Appl
3	60	12.0	5452 2	US-09-130-114-1 Sequence 1, Appl
4	60	12.0	9600 4	US-08-910-647-1 Sequence 1, Appl
5	60	12.0	10596 1	US-07-884-811-15 Sequence 15, Appl
6	60	12.0	10596 1	US-07-885-971-15 Sequence 15, Appl
7	60	12.0	10596 1	US-08-087-763A-15 Sequence 15, Appl
8	60	12.0	10596 1	US-08-194-088B-15 Sequence 15, Appl
9	60	12.0	10596 2	US-08-194-088B-15 Sequence 15, Appl
10	60	12.0	10596 5	PCT-US93-0468-15 Sequence 15, Appl
11	57.2	11.4	15333 3	US-09-128-153-16 Sequence 16, Appl
12	57.2	11.4	313 3	US-09-165-264-8 Sequence 8, Appl
13	50	10.0	320 4	US-09-165-264-13 Sequence 13, Appl
14	49.2	9.8	3661 4	US-08-718-388-5 Sequence 5, Appl
15	49.2	9.8	7824 4	US-08-718-388-6 Sequence 6, Appl
16	49.2	9.8	16382 4	US-08-718-388-8 Sequence 8, Appl
17	48.8	9.7	4411529 4	US-09-103-840A-1 Sequence 1, Appl
18	48.4	9.7	3000 4	US-08-460-265C-5 Sequence 5, Appl
19	48	9.6	7218 1	US-08-332-463-14 Sequence 14, Appl
20	47.8	9.6	320 4	US-09-165-264-7 Sequence 7, Appl
21	47	9.4	44337 2	US-08-804-227C-7 Sequence 7, Appl
22	46.6	9.3	2793 1	US-08-804-198-1 Sequence 1, Appl
23	46.6	9.3	2793 1	US-08-209-747-1 Sequence 1, Appl
24	46.6	9.3	2793 1	US-08-458-298-1 Sequence 1, Appl
25	46.4	9.2	3069 3	US-08-335-865J-7 Sequence 7, Appl
26	45.8	9.2	320 4	US-09-165-264-11 Sequence 11, Appl
27	45.8	9.2	16885 1	US-08-390-878-16 Sequence 16, Appl

28	45.4	9.1	1028 4	US-08-118-200-1 Sequence 1, Appl
29	45.4	9.1	1028 4	US-08-458-745-1 Sequence 1, Appl
30	45.2	9.0	320 4	US-09-165-264-14 Sequence 14, Appl
31	45	9.0	318 4	US-09-165-264-12 Sequence 12, Appl
32	45	9.0	8931 3	US-09-028-934-28 Sequence 28, Appl
33	44.8	8.9	4403765 4	US-09-103-840A-2 Sequence 2, Appl
34	44.4	8.9	8438 1	US-07-945-283-1 Sequence 1, Appl
35	44.2	8.8	4403765 4	US-09-103-840A-2 Sequence 2, Appl
36	44	8.8	4411529 4	US-09-103-840A-1 Sequence 1, Appl
37	43.8	8.8	3147 2	US-08-781-802-7 Sequence 7, Appl
38	43.8	8.8	3147 2	US-08-694-878-7 Sequence 7, Appl
39	43.8	8.8	3147 2	US-09-058-260-7 Sequence 7, Appl
40	43.6	8.7	2150 2	US-08-318-837-1 Sequence 1, Appl
41	43.6	8.7	6530 2	US-08-146-930-1 Sequence 1, Appl
42	43.6	8.7	6530 3	US-08-458-240-1 Sequence 1, Appl
43	43.6	8.7	6530 5	PCT-US93-03993-1 Sequence 1, Appl
44	43.4	8.7	2150 3	US-09-263-023-1 Sequence 1, Appl
45	43.4	8.7	3472 6	5244792-1 Patent No. 5244792

ALIGNMENTS

RESULT 1
US-09-050-863-2/c
Sequence 2, Application US/09050863
Patent No. 6114111
GENERAL INFORMATION:
APPLICANT: Lao, Ying
APPLICANT: Hwang, Betty
APPLICANT: Payer, Don
TITLE OF INVENTION: Mammalian Protein Interaction Cloning
TITLE OF INVENTION: System
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hoback, Test, Albritton & Herbert
STREET: 4 Embarradero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050, 863
FILING DATE: 30-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJR/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 949-8711
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-09-050-863-2
Query Match 12.0%; Score 60; DB 3; Length 2580;
Best Local Similarity 50.6%; Pred. No. 0.00034;
Matches 170; Conservative 0; Mismatches 165; Indels 1; Gaps 1;
QY 131 GCCCAGATAGAGCGCCGCGGCGGCTGCTGAGATCCCGCGCGCTGCTGCGCGCGG 190
DB 1021 GCCCTCTGCGCCCTCTGCGCCCTCTGCTGCTGCGCCCTCTGCTGCTGCTGCTGCT 962

OY	191	AGAGCCCAACCCCGCCGCGCGCGCGCATCTTGGACAGTGGAGGCGCGGCACATGAGAGGATTA	250
Db	961	GGCCCTCTGACCCCGCCCTCTGCTCTGAGCCCGCTGAGCCCGCTCTGCTGCTCTGAGCCCGCT	902
OY	251	CCGCGACATGAAAGACCCCGCGCGGCGAGCGCTGTACAGCGCGCGCCGCGCTGACAG	310
Db	901	CGTCTCTCTCTGACCGCTCTCTGACCGCTCTGACCGCGCTCTCTGCTCTGCTCTCTCTCTCTGCG	84.2
OY	311	CGCGATCTGCG-ACTGAGCCGCTACTGAGAAACACTCTTCTGCGCGCGCGCATATCTCGCTTCT	369
Db	841	CGTCTCTCTGCTCTGAGCCCGCTCTCTCTGCTCTGCTCTGCGCGCTCTGCGCGCTCTCTCTCT	78.2
OY	370	CGACGCGACAGCGCGCGCTTCAATGAGCGCGCGCGCTTCTCTCTGAGAGGACAGCGCGCGCAG	429
Db	781	CGTCTCTCTGCGCGCGCTCTCTCTGCTCTGCTCTGCGCGCGCTCTCTGCGCGCTCTCTCTCTGAGCCCT	72.2
OY	430	ACACACAGCATGCGCGCGCGGCGGAGCGACAGCGCTCTGACCT	465
Db	721	CGTCTCTCTCTGCGCGCGCTCTCTCTGCTCTGCTCTCTGCGCGCT	686

```

1      RESULT 2
2      US-09-359-081-2/C
3      : Sequence 2, Application US/09359081
4      : Patent No. 6316223
5      : GENERAL INFORMATION:
6      : APPLICANT: Lao, Ying
7      :           Hiang, Betty
8      :           Payan, Don
9      : TITLE OF INVENTION: Mammalian Protein Interaction Cloning
10     : System
11     : NUMBER OF SEQUENCES: 5
12     : CORRESPONDENCE ADDRESS:
13     : ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
14     : STREET: 4 Embarradero Center, Suite 3400
15     : CITY: San Francisco
16     : STATE: CA
17     : COUNTRY: USA
18     : ZIP: 94111-4187
19     : COMPUTER READABLE FORM:
20     : MEDIUM TYPE: Floppy disk
21     : COMPUTER: IBM PC compatible
22     : OPERATING SYSTEM: PC DOS/MS-DOS
23     : SOFTWARE: PatentIn Release #1.0, Version #1.30
24     : CURRENT APPLICATION DATA:
25     : APPLICATION NUMBER: US/09/359,081
26     : FILING DATE: 22-Jul-1999
27     : CLASSIFICATION: <Unknown>
28     : PRIOR APPLICATION DATA:
29     : APPLICATION NUMBER: 09/050,863
30     : FILING DATE: <Unknown>
31     : ATTORNEY/AGENT INFORMATION:
32     : NAME: Silva, Robin M.
33     : REGISTRATION NUMBER: 38,304
34     : REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
35     : TELECOMMUNICATION INFORMATION:
36     : TELEPHONE: (415) 781-1989
37     : TELEFAX: (415) 949-8711
38     : INFORMATION FOR SEQ ID NO: 2:
39     : SEQUENCE CHARACTERISTICS:
40     : LENGTH: 2580 base pairs
41     : TYPE: nucleic acid
42     : STRANDEDNESS: unknown
43     : TOPOLOGY: unknown
44     : MOLECULE TYPE: DNA
45     : SEQUENCE DESCRIPTION: SEQ ID NO: 2:
46     :
47     :
48     :
49     :
50     :
51     :
52     :
53     :
54     :
55     :
56     :
57     :
58     :
59     :
60     :
61     :
62     :
63     :
64     :
65     :
66     :
67     :
68     :
69     :
70     :
71     :
72     :
73     :
74     :
75     :
76     :
77     :
78     :
79     :
80     :
81     :
82     :
83     :
84     :
85     :
86     :
87     :
88     :
89     :
90     :
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :
622    :
623    :
624    :
625    :
626    :
627    :
628    :
629    :
630    :
631    :
632    :
633    :
634
```

Qy	131	GGCAGATGACGCGCCCGGCGGCGCCCGGGATGCGAGATGCCGCGCTGATGCGCGCG	190
Db	1021	GGCCCTCGGCCCTCTGACCCGCTGACGCTGACCCCTCTCTCGCTCTGCTGCTGCTCTCT	962
Qy	191	AGGAGCCGACCCGCGCTGGCGGCGGATGCTGGACGAGGGGCGCCGCAATGGACGAGATTA	250
Db	961	GGCCCTCGGCCCTCTCTCTGCTGCGCCGCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCT	902
Qy	251	CCGCGAGATGAACCCCGCGCGGCGAGCCTCTGATACAGCCGCGCGCCCTGACAG	310
Db	901	CCCTCGCTGCTCGCCCTCTCTGACCCCTCTGCGCCCTCTCTGCTCTGCTGCGCCCTCTGCTC	842
Qy	311	CGCAGATGCTC-ACTGACCCCTACTGGAACACTCTTCGCGTGCACGACATACCGCTCTCT	369
Db	841	CCGCTCTGCTGCTGCTGCGCCCTCTCTCTGCTGCTGCGCCCTCTCTGCGCCCTCTCT	782
Qy	370	CCAGGAGATGCGCGCGCTTCATGAGCTGCGGCTCTCTCTCGCGAGCGACGCGCCGAG	429
Db	781	CGAGCTCTGCGCCCTCTCTCTGCTGCTGCGCCCTCTCTGCGCCCTCTCTGCGCCCTCTCT	722
Qy	430	ACAGAGCTATGCGCCCGCGAGCGACGCGCTGACGCT	465
Db	721	CCGCTGCTGCTGCGCCCTCTCTGCTGCTGCGCCCT	686

[illegible]

US-08-194-088B-15/c
Sequence 15, Application US/08194088B

Patent No. 5580963
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,088B
FILING DATE: 09-FEB-1994
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gallegos, R. Thomas
REGISTRATION NUMBER: 32,692
REFERENCE/DOCKET NUMBER: 755D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-2614
TELEFAX: 415/952-9881

TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-194-088B-15

Query Match 12.0%; Score 60; DB 1; Length 10596;
Best Local Similarity 50.6%; Pred. No. 0.00037;
Matches 170; Conservative 0; Mismatches 165; Indels 1; Gaps 1;

QY 131 GCCCAGATAGACGCCGCCGCGGCGGCTCTGAGATCCCGCGCTGCTGCGCGCG 190
DB 2601 GCCCTCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2542
QY 191 AGGACCCGACCG 250
DB 2541 GCCCTCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482
QY 251 CCCGACGATGACGCCGCCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 310
DB 2481 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2422
QY 311 CGCAGATGACG 369
DB 2421 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2362
QY 370 CAGGACGACGCCGCCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 429
DB 2361 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2302
QY 430 ACACGAGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
DB 2301 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2266

RESULT 9
US-08-194-087-15/c

Sequence 15, Application US/08194087
Patent No. 5879910

GENERAL INFORMATION:
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,087
FILING DATE: 18-MAY-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 779
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881

TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-194-087-15

Query Match 12.0%; Score 60; DB 2; Length 10596;
Best Local Similarity 50.6%; Pred. No. 0.00037;
Matches 170; Conservative 0; Mismatches 165; Indels 1; Gaps 1;

QY 131 GCCCAGATAGACGCCGCCGCGGCGGCTCTGAGATCCCGCGCTGCTGCGCGCG 190
DB 2601 GCCCTCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2542
QY 191 AGGACCCGACCG 250
DB 2541 GCCCTCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482
QY 251 CCCGACGATGACGCCGCCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 310
DB 2481 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2422
QY 311 CGCAGATGACG 369
DB 2421 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2362
QY 370 CAGGACGACGCCGCCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 429
DB 2361 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2302
QY 430 ACACGAGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 465
DB 2301 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2266

RESULT 10
PCT-US93-04648-15/c
Sequence 15, Application PC/TUS9304648


```
SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-8
```

```
Query Match          10.4%; Score 52.2; DB 4; Length 319;
Best Local Similarity 48.2%; Pred. No. 0.011;
Matches 147; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
```

```
OY 157 GGGTCTGAGATCCCGCGCTGCTGCGCGGCGAGACCCCGCGCGCGCGCA 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 GGGTCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 253
OY 217 TGCCTGCAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 252 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 193
OY 277 GCAAGCTCTGTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133
OY 337 ACAACCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73
OY 397 CCGCGCTCTCTCTGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13
OY 457 CCTCC 461
    |||||
Db 12 CCGCG 8
```

```
RESULT 13
US-09-165-264-13/c
; Sequence 13, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuralayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-13
```

```
Query Match          10.0%; Score 50; DB 4; Length 320;
Best Local Similarity 47.5%; Pred. No. 0.029;
Matches 149; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
```

```
OY 148 GCGCGCGCGCGCTGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 GCGATGCTCTGTCAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 260
OY 208 GCGCGCGCGCGCTGAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 200
OY 268 GCGCGCGCGCGCGCTGATACCGCGCGCGCGCGCGCGCGCGCGCGCG 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 140
```

```
OY 328 CTTACTGAGACACCTTCTGCTGCCGCCATACCTGCTTCCAGCAGACCGCGCT 387
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 80
OY 388 TCATGAGCTCGCGCTCTGCTGCGAGCGCGCGCGCGCGCGCGCGCGCG 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20
OY 448 TGGCGACCGCTCC 461
    |||||
Db 19 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6
```

```
RESULT 14
US-08-718-388-5/c
; Sequence 5, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:
; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING Igg Fc REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,388
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3661 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-718-388-5
```

```
Query Match          9.8%; Score 49.2; DB 4; Length 3661;
Best Local Similarity 45.9%; Pred. No. 0.049;
Matches 168; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
```

```
OY 54 CCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2805 CAGAGGCTCTCAGATGCGCTGTACTTGCGCAGCGCTCGCAGCAGCGCTGGAATAC 2746
OY 114 CCACGACGAGGAGAGAACCCAGATAGACGCCCGCGCGCGCGCGCGCGCG 173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2745 TGCAGGCGGCGCACAAAGCGCGTGGGAGGGGCGCAGGGGCGCGGTGGGAGATCAG 2686
OY 174 CCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2685 CCGCAGGGGCTCGGGGCGCGCGAGGAGCTTGTGCTGTGGGTGACAGCGGCGCAT 2626
OY 234 GCCATGACAGGAGATTACCGCGAGCATGAACCGCGCGCGCGCGCGCGCGCG 293
```


Db 2625 GGGTTGACACACATTCCTCCAGCCCTGGGGGCGCCACCTCCATCCGGGGGCTTC 2566
OY 294 CCGCCGCGCCCTGACAGAGCCCATGCTGCACTGCCCTACTGGAACACTTCTGCTGCCG 353
Db 2565 CCGCCACCGCGCTTCAGTGTCTCTGCGGGGCTCTGTTGTAGTTCGCGATAAGCCACAG 2506
OY 354 CCATACCCCTGCTTCTCTCCAGGACAGCGCCCGCTTCATGAGTCCGCCCTCTCTCGGC 413
Db 2505 AGAGAGCCCGGTAAGCGCGCGACAGCGCAGCGACGAGCTGTCCCGTGAAGGCC 2446
OY 414 AGCCAG 419
Db 2445 AGCGAG 2440

RESULT 15

US-08-718-388-6/c
Sequence 6, Application US/08718388

Patent No. 6271362

GENERAL INFORMATION:

APPLICANT: MORIKAWA, MINORU

APPLICANT: HARADA, NAOKI

TITLE OF INVENTION: GENE ENCODING Igg Fc REGION-BINDING

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

STREET: PO BOX 747

CITY: FALLS CHURCH

STATE: VA

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/718,388

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: MURPHY JR, GERALD M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 0230-111

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 7824 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 21..7802

US-08-718-388-6

Query Match

Best Local Similarity 45.9%; Score 49.2; DB 4; Length 7824;

Matches 168; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

OY 54 CCGCGGCTCCCGCGCCCGCGAGCCCGCGAGATCATGATGCTCCGCCACCGCGC 113

Db 6940 CAGAGGCTCTCCAGATGGCTTGAACCTTGGAGGCTCAGCAAGAGCTGAGATAC 6881

OY 114 CCACACGAG 173

Db 6880 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6821

OY 174 GCGCTGTGCCCGCGAGAGACCCCGCCCTGCGCCCGGATGCTTGCATGAGGGGCC 233
Db 6820 CCGAGGGGCTCCGGGCGCGGAGAGACTTCTGCTGTGGGTCACAGGCGAGGGCAT 6761
OY 234 GCCATGACAGGAGATTACCGCAGCATGAAACCCCGCGCGGCGGCGGCGGCGGCGGCGG 293
Db 6760 GCGTTGACACACATTCCTCCGCGAGCCCTGGCGCGCGCCACTGCGCATTCGGGGGCTTC 6701
OY 294 CCGCGGCGCTGCAGAGCGCATGCTGCACTGCGCCCTACTGGAACACTTCTGCTGCCG 353
Db 6700 CCGCCACCTGCTTCAGTGTCTGTGGGGGCTCCTGTTGTAGTTCGCGCATAAAGCCACAG 6641
OY 354 CCATACCCCTGCTTCTTCAAGCGAGAGCGCGCGCTTCATGAGCTCCGCTCTCTGCGGC 413
Db 6640 AGAGAGCGCGGTAAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6581
OY 414 AGCCAG 419
Db 6580 AGCGAG 6575

Search completed: October 10, 2002, 19:27:34
Job time : 104.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 695.733 Seconds

(without alignments)
9699,805 Million cell updates/sec

Title: US-09-489-101a-13_COPY_1_500

Sequence: 1 gccgcaaacccggaagtgag.....gcgacttgcgcagactcc 500

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : EST:*

1: em_estha:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	460	92.0	483	10	BM148224 TCAP1098
2	396.2	79.2	402	10	BE241632 TCAP1098
3	178	35.6	712	10	BE409724 TCAP1098
4	121	24.2	635	9	BE664186 BE664186
5	116.2	23.2	695	10	BE952238 BE952238
6	112.8	22.6	131	10	BE242275 TCAP1098
7	112.8	22.6	2658	11	BC019206 Mus muscu
8	106.6	21.3	635	12	AG084012 Pan trogl
9	79.2	15.8	776	12	CNS0109Y
10	78.4	15.7	951	9	AL543362
11	77.4	15.5	612	10	BE149078
12	75	15.0	274	9	AA948370
13	74.8	15.0	834	10	BE141657
14	73.2	14.6	855	12	AG030581
15	72.8	14.6	798	10	BE141695
16	72.8	14.6	848	12	CNS020YE
17	72.4	14.5	559	9	AL579455

C 18	72.4	14.5	932	12	CNS0072Q	AL066742 Drosophila
C 19	71.6	14.3	810	12	AG060267	Pan trogl
C 20	71.1	14.2	741	12	A2194927	SP_1029_A
C 21	70.6	14.1	576	10	BE786264	SEAMC006
C 22	70.4	14.1	927	10	BE786316	SEAMC006
C 23	70.2	14.0	1139	12	AG137991	Pan trogl
C 24	70.0	14.0	925	12	CNS0091P	AG1053013 Drosophila
C 25	70.0	14.0	1023	12	AG128304	Pan trogl
C 26	70.0	14.0	1101	12	CNS0153F	AG104949 Drosophila
C 27	69.8	14.0	918	10	BE809598	BE809598 mgct001xb
C 28	69.6	13.9	782	10	BE809582	BE809582 mgct001xa
C 29	69.6	13.9	831	10	BE952198	BE952198 HYSM000
C 30	69.6	13.9	1080	12	AG093056	AG093056 HS_4832_A
C 31	69.4	13.9	821	12	AG093474	AG093474 HS_4832_A
C 32	69.4	13.9	1128	10	BM477735	BM477735 ACENCOU
C 33	69.2	13.8	641	9	AL357868	qv13b02.x
C 34	69.2	13.8	925	10	BE441241	BE441241 GA_Ea001
C 35	69.0	13.8	908	12	CNS006B4	AL064031 Drosophila
C 36	68.6	13.7	643	12	AG123077	AG123077 Pan trogl
C 37	68.4	13.7	827	12	AG183795	AG183795 SP_1002_A
C 38	68.2	13.6	613	12	AG043036	AG043036 Pan trogl
C 39	68.2	13.6	670	12	AG126228	AG126228 Pan trogl
C 40	68.2	13.6	836	12	AG185303	AG185303 SP_1005_A
C 41	68.2	13.6	863	12	AG132157	AG132157 Pan trogl
C 42	68.2	13.6	864	12	AG060559	AG060559 Pan trogl
C 43	68.2	13.6	880	12	AG074253	AG074253 HS_5383_B
C 44	68.2	13.6	925	9	AL581447	AL581447
C 45	68.2	13.6	925	12	CNS0091P	AL053013 Drosophila

ALIGNMENTS

RESULT 1
BM148224 483 bp mRNA linear EST 30-NOV-2001
LOCUS TCAP109825 Pediatric acute myelogenous leukemia cell (FAB M1)
DEFINITION Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAP109825, mRNA sequence.

ACCESSION BM148224.1 GI:17168744
VERSION BM148224
KEYWORDS EST.
SOURCE Human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 483)
AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R., Jr., Gunaratne,P.H., Mizuy,D., Bouck,J., Gibbs,R.A., and Margolin,J.F.
JOURNAL Pediatric Leukemia cDNA Sequencing Project (2001)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@xccc.org
Seq primer: M13 primer:
Location/Qualifiers

FEATURES
source
1..483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAP109825"
/clone_lib="Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10B"
/note="Vector: lambda PSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored

BASE COUNT 121 a 204 c 182 g 126 t 2 others
 ORIGIN
 Query Match 24.2% Score 121; DB 9; Length 635;
 Best Local Similarity 60.48; Pred No. 4.4e-11;
 Matches 299; Conservative 0; Mismatches 145; Indels 51; Gaps 4;

1 GCCGAAACCCGAGAGTGAAGGCGGCGAGCTGCGAGCTCGAGAAACGCGCGG 60
 6 GCCGAAACCCGAGAGTGAAGGCGGCGAGCTGCGAGAGTGAAGGCGCGG 63
 61 CTCGCGCGCGGCGAGAGTGAAGGCGGCGAGCTGCGAGAGTGAAGGCGG 120
 64 --GGGCGAGAGTGAAGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 120
 121 GGAGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 180
 121 AGAGTGAAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 175
 181 TGCCGCGCGGAGAGTGAAGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 240
 176 TGCCGCGCGGAGAGTGAAGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 206
 241 ACAGGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGGCGG 300
 207 ATGGGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGGCGG 256
 301 CCTGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 360
 257 --CTACGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 314
 361 CTCGCTTCTCAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 420
 315 CCAGCTTCTCAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 374
 421 CCTGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 480
 375 CTGCGCTTCTCAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 434
 481 GCGAGTGAAGGCGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 495
 435 GCGAGTGAAGGCGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 449

RESULT 5
 BE952238
 LOCUS
 DEFINITION
 BE952238
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT

BE952238 695 bp mRNA linear EST 04-OCT-2000
 UI-M-CD0-ayk-d-03-0-UI-s1 NIH BMAP Ret2 Mus musculus cDNA clone
 BE952238
 BE952238.1 GI:10592495
 EST.
 house mouse.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 695)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20899-3643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: meste@nimh.nih.gov
 Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares

is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 Seq. Primer: M13 Forward
 POLY-A-MO. Location/Qualifiers

1.695
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CD0-ayk-d-03-0-UI"
 /clone_id="M13_BMAP_Ret2"
 /dev_stage="1 day"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP Ret2 library is derived from mouse neonatal
 retina tissue. For a detailed description of the library
 from which this clone was derived, please visit our web
 site at brainest.eng.uiowa.edu.
 TAG-Seq=None found"
 BASE COUNT 145 a 234 c 173 g 143 t

Query Match 23.2% Score 116.2; DB 10; Length 695;
 Best Local Similarity 64.9% Pred. No. 2.8e-10;
 Matches 194; Conservative 0; Mismatches 93; Indels 12; Gaps 1;

202 CCGCGCGCGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 261
 140 CTCGAGTGAAGGCGGCGGCGGCGAGCTGCGAGAGTGAAGGCGGCGG 199
 262 AACGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 321
 200 AGCTGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 247
 322 ACTGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 381
 248 ACTGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 307
 382 GCGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 441
 308 AACGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 367
 442 CCGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 500
 368 CCGCGCGCGGCGAGAGTGAAGGCGGCGGCGAGCTGCGAGAGTGAAGGCGG 426

RESULT 6
 BE242275
 LOCUS
 DEFINITION
 BE242275
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT

BE242275 131 bp mRNA linear EST 03-OCT-2001
 TCAAP1545 Pediatric acute myelogenous leukemia cell (FAB M1)
 Baylor-HGSC project-TCAAP1545, mRNA
 BE242275
 BE242275.1 GI:9094002
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 131)
 Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, J.R., Muzny, D.
 Bouck, J., Gibbs, R.A. and Margolin, J.F.
 Pediatric leukemia cDNA sequencing project
 Unpublished (2000)
 Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038

Email: clones@cycc.org
 Curation: Carninci, P. and Hayashizaki, Y. High efficiency
 full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Seq primer: M13 primer.

FEATURES

Location/Qualifiers
 1..131
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCAP1545"
 /clone_1lb="pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HESC project-TCAP"
 /sex="male"
 /tissue_type="leukopheresis"
 /cell_type="myeloid cell"
 /dev_stage="pediatric 6 years"
 /lab_host="DH10B"
 /note="Vector: lambda psb. Site_1: BamHI; Site_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer [5'GAGAGCTGAGCGCGCGAGAGAG(T)VN 3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand was primed with a BamHI-dC primer [5'AGAGAGCTGAGCGCGCGCGCGCAATTAATTAAT(C) 3']. Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and SalI sites of lambda psb vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P., Westover A., Nishiyama Y., Ohsumi T., Itoh M., Nagaoka S., Sasaki, Okazaki Y., Muramatsu M., Schneider C., Hayashizaki Y., High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)".

BASE COUNT 19 a 47 c 49 g 15 t 1 others
 ORIGIN

Query Match 22.6% Score 112.8; DB 10; Length 131;
 Best Local Similarity 93.6% Pred. No. 9.3e-10;
 Matches 117; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 22 GCGCGAGCTGCGAGCTGCGAGAAAGCGCGCGCTCGCGCGCGAGCCG 81
 |||||
 Db 2 GCGCGAGCTGCGAGCTGCGAGAAAGCGCGCGCTCGCGCGCGAGCCG 61
 OY 82 GCGCGAGCTGCGAGCTGCGAGCGCGCGCGCGCGCGCGAGCGAGCGAG 141
 |||||
 Db 62 GCGCGAGCTGCGAGCTGCGAGCGCGCGCGCGCGCGCGAGCGAGCGAG 121
 OY 142 CGGCC 146
 |||||
 Db 122 CGGCC 126

RESULT 7
 LOCUS BC019206 2658 bp mRNA linear HFC 11-DEC-2001
 DEFINITION Mus musculus, similar to KIAA0963 protein, clone IMAGE:5052084,
 mRNA.

ACCESSION BC019206
 VERSION BC019206.1 GI:17512508
 KEYWORDS HFC.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2658)
 AUTHORS Strausberg R.
 TITLE Direct Substitution
 JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530,
 USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nih.gov

Shcherbako, Y., Wetherby, R.D., Beckstrom-Sternberg, S.M.,
 Benjamini, E., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, O.L., Mastaglio, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantripo, S., Thomas, P.J.,
 Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: FRK Plate: 36 Row: m Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein
 This clone has the following problem: incomplete processing.

FEATURES

Location/Qualifiers
 1..2658

source
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5052084"
 /tissue_type="Liver, normal, 5 month old male mouse."
 /clone_1lb="M1 CGAP_119"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"

BASE COUNT 600 a 761 c 696 g 601 t
 ORIGIN

Query Match 22.6% Score 112.8; DB 11; Length 2658;
 Best Local Similarity 64.6% Pred. No. 1.1e-09;
 Matches 190; Conservative 0; Mismatches 92; Indels 12; Gaps 1;

OY 202 CGCGCTGCGCGAGCTGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 261
 |||||
 Db 139 CGCGCTGCGCGAGCTGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 198
 OY 262 AACCCCGCGCGCGAGCTGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 321
 |||||
 Db 199 AGCTCCACCGCGCGAGCTGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 246
 OY 322 ACTGCCCTACTGGAACACTTCTGCTGCGCGCGAGTGGAGTGAACGAGTGA 381
 |||||
 Db 247 ACTGCCCTACTGGAACACTTCTGCTGCGCGCGAGTGGAGTGAACGAGTGA 306
 OY 382 GCGCGTCAAGAGCTGCGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 441
 |||||
 Db 307 AACAGTTCGCGGCTTCCAGCCATCTTGGTGGCGAGTGGAGTGAACGAGTGA 366
 OY 442 CCCCCGCGCGCGAGCTGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 495
 |||||
 Db 367 CCACACAGCGCGCGAGCTGCGAGTGGGCGCGCGAGTGGAGTGAACGAGTGA 420

RESULT 8
 LOCUS AG084012 635 bp DNA linear GSS 03-NOV-2001
 DEFINITION Pan troglodytes DNA, clone: PTB-081K02.R, genomic survey sequence.
 ACCESSION AG084012
 VERSION AG084012.1 GI:16635814

KEYWORDS GSS: GSS (genome survey sequence).
 SOURCE Pan troglodytes male lymphoblast DNA, clone: PTB-081K02.R.
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (sites)
 Email: cgabs-r@mail.nih.gov

AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.			
TITLE	BAC end sequences of library PTB			
JOURNAL	Unpublished			
REFERENCE	2 (Bases 1 to 635)			
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-AUG-2001) Aao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-chou, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan			
REFERENCE	tel:81-45-503-9110, fax:81-45-503-9170			
AUTHORS	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.			
COMMENT	PRIMERS			
	Sequencing: M13Rev			
	LIBRARY			
	Vector : pKS145			
	R.site 1 : SacI			
	R.site 2 : SacI.			
FEATURES	Location/Qualifiers			
source	1..635			
	/organism="Pan troglodytes"			
	/db_xref="taxon:9598"			
	/clone="PTB-081K02.R"			
	/sex="male"			
	/cell_type="lymphoblast"			
	/clone_lib="PTB chimpanzee Male BAC Library"			
BASE COUNT	104 a	236 c	189 g	105 t
ORIGIN	1 others			
Query Match	21.3%	Score 106.6;	DB 12;	Length 635;
Best Local Similarity	95.6%	Pred. No. 1,1e-08;		
Matches 109;	Conservative	0;	Mismatches	5;
			Indels	0;
			Gaps	0;
OY	382	GGCCGTCATGACATCCGCGCTCTCTCTCTGCGACACACCTGCGCCAGACAGCTTANG	441	
Db	23	GCNCTCTCGAGACCTCGCGCTCTCTCTCTGCGAGCACCTCTGCGACACACCTTANG	82	
OY	442	CCCCCGTGGCCACCGCGCTCTCTCTCTGCGACACCAAGACCTGCGACTTCTCTGAG	495	
Db	83	CCCCCGTGGCCACCGCGCTCTCTCTCTGCGACACCAAGACCTGCGACTTCTCTGAG	136	
RESULT 9				
CNS010RY				
LOCUS	CNS010RY 776 bp DNA linear GSS 26-JUL-1999			
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC			
LOCATION	BACN04P24 of DrosBAC library from Drosophila melanogaster (fruit			
ACCESSION	AL099352			
VERSION	AL099352.1 GI:5610963			
KEYWORDS	GSS.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (Bases 1 to 776)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage :			
COMMENT	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr			
	- Web : www.genoscope.cns.fr)			
	determination of this BAC-end sequence was carried out as part of a			
	collaboration with the European Drosophila genome project (BDGP) -			
	http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC			
	library (Dros BAC) was made by Alain Billaud at CERN (Centre			
	d'Etude du Polymorphisme Humain) with funding provided by a IRC			
	Project grant. The DNA was prepared from embryos by Alain Bucheton			
	and Genevieve Payan. It has been constructed in the vector			

[illegible]

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrived by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: t1AM114 row: m column: 13
 High quality sequence start: 14
 High quality sequence stop: 581.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
(bases 1 to 274)				
NCI-CGAP	http://www.ncbi.nlm.nih.gov/hc/cgap.			
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
Tumor Gene Index				
Unpublished (1997)				
Contact: Robert Strausberg, Ph.D.				
Email: cgapbs-r@mail.nih.gov				
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.				
Emmert-Buck, M.D., Ph.D.				
cDNA Library Preparation: M. Bento Soares, Ph.D.				
cDNA Library Arrayed by: Greg Lennon, Ph.D.				
DNA Sequencing by: Washington University Genome Sequencing Center				
clone distribution: NCI-CGAP clone distribution information can be				
found through the I.M.A.G.E. Consortium/BLNL at:				
www.bio.lim.gov/dbfp/image/image.html				
Seq primer: -40m13 fwd. ET from Amersham				
High quality sequence stop: 228.				
Location/Qualifiers				

```
FEATURES      location/Qualifiers
source        1. 274
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
```

Query Match	15.0%;	Score 74.8;	DB 10;	Length 834;
Best Local Similarity	44.4%;	Pred. No. 0.0021;		

/organism="Pan troglodytes"

PR 14-AUG-2000; 2000US-0225267.

PR	14-ANG-2000	2000US-02255.66
PR	14-ANG-2000	2000US-02255.70
PR	14-ANG-2000	2000US-02255.77
PR	14-ANG-2000	2000US-02255.84
PR	14-ANG-2000	2000US-02257.58
PR	14-ANG-2000	2000US-02257.59
PR	18-ANG-2000	2000US-02263.79
PR	22-ANG-2000	2000US-02266.81
PR	22-ANG-2000	2000US-02266.88
PR	22-ANG-2000	2000US-02271.82
PR	30-ANG-2000	2000US-02285.24
PR	01-SEP-2000	2000US-02292.87
PR	01-SEP-2000	2000US-02293.43
PR	01-SEP-2000	2000US-02293.44
PR	01-SEP-2000	2000US-02293.45
PR	05-SEP-2000	2000US-02295.09
PR	05-SEP-2000	2000US-02295.13
PR	05-SEP-2000	2000US-02304.37
PR	06-SEP-2000	2000US-02304.38
PR	08-SEP-2000	2000US-02313.42
PR	08-SEP-2000	2000US-02313.43
PR	08-SEP-2000	2000US-02314.43
PR	08-SEP-2000	2000US-02314.14
PR	08-SEP-2000	2000US-02316.80
PR	12-SEP-2000	2000US-02320.81
PR	12-SEP-2000	2000US-02321.68
PR	14-SEP-2000	2000US-02323.97
PR	14-SEP-2000	2000US-02323.98
PR	14-SEP-2000	2000US-02323.99
PR	14-SEP-2000	2000US-02324.00
PR	14-SEP-2000	2000US-02324.01
PR	14-SEP-2000	2000US-02330.63
PR	14-SEP-2000	2000US-02330.64
PR	14-SEP-2000	2000US-02330.65
PR	21-SEP-2000	2000US-02343.72
PR	21-SEP-2000	2000US-02343.73
PR	25-SEP-2000	2000US-02345.97
PR	25-SEP-2000	2000US-02345.98
PR	25-SEP-2000	2000US-02354.84
PR	27-SEP-2000	2000US-02358.34
PR	27-SEP-2000	2000US-02358.36
PR	29-SEP-2000	2000US-02361.27
PR	29-SEP-2000	2000US-02361.27
PR	29-SEP-2000	2000US-02363.68
PR	29-SEP-2000	2000US-02363.69
PR	29-SEP-2000	2000US-02363.70
PR	02-OCT-2000	2000US-02365.02
PR	02-OCT-2000	2000US-02370.37
PR	02-OCT-2000	2000US-02370.38
PR	02-OCT-2000	2000US-02370.39
PR	02-OCT-2000	2000US-02370.40
PR	02-OCT-2000	2000US-02373.73
PR	02-OCT-2000	2000US-02373.74
PR	02-OCT-2000	2000US-02373.75
PR	13-OCT-2000	2000US-02396.37
PR	13-OCT-2000	2000US-02396.37
PR	20-OCT-2000	2000US-02409.60
PR	20-OCT-2000	2000US-02412.61
PR	20-OCT-2000	2000US-02417.85
PR	20-OCT-2000	2000US-02417.86
PR	20-OCT-2000	2000US-02417.87
PR	20-OCT-2000	2000US-02418.08
PR	20-OCT-2000	2000US-02418.09
PR	20-OCT-2000	2000US-02418.16
PR	01-NOV-2000	2000US-02446.17
PR	08-NOV-2000	2000US-02464.74
PR	08-NOV-2000	2000US-02464.75
PR	08-NOV-2000	2000US-02464.76
PR	08-NOV-2000	2000US-02464.77
PR	08-NOV-2000	2000US-02464.78
PR	08-NOV-2000	2000US-02465.23
PR	08-NOV-2000	2000US-02465.24
PR	08-NOV-2000	2000US-02465.25
PR	08-NOV-2000	2000US-02465.26

08-NOV-2000; 2000US-0246527.
 08-NOV-2000; 2000US-0246528.
 08-NOV-2000; 2000US-0246532.
 08-NOV-2000; 2000US-0246609.
 08-NOV-2000; 2000US-0246610.
 08-NOV-2000; 2000US-0246611.
 08-NOV-2000; 2000US-0246613.
 17-NOV-2000; 2000US-0249207.
 17-NOV-2000; 2000US-0249208.
 17-NOV-2000; 2000US-0249209.
 17-NOV-2000; 2000US-0249210.
 17-NOV-2000; 2000US-0249211.
 17-NOV-2000; 2000US-0249212.
 17-NOV-2000; 2000US-0249213.
 17-NOV-2000; 2000US-0249214.
 17-NOV-2000; 2000US-0249215.
 17-NOV-2000; 2000US-0249216.
 17-NOV-2000; 2000US-0249217.
 17-NOV-2000; 2000US-0249218.
 17-NOV-2000; 2000US-0249244.
 17-NOV-2000; 2000US-0249245.
 17-NOV-2000; 2000US-0249265.
 17-NOV-2000; 2000US-0249297.
 17-NOV-2000; 2000US-0249299.
 17-NOV-2000; 2000US-0249300.
 01-DEC-2000; 2000US-0250160.
 01-DEC-2000; 2000US-0250391.
 05-DEC-2000; 2000US-0251030.
 05-DEC-2000; 2000US-0251988.
 05-DEC-2000; 2000US-0256719.
 06-DEC-2000; 2000US-0251479.
 08-DEC-2000; 2000US-0251856.
 08-DEC-2000; 2000US-0251868.
 08-DEC-2000; 2000US-0251869.
 08-DEC-2000; 2000US-0251989.
 11-DEC-2000; 2000US-0251990.
 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI: 2001-451937/48.
 Isolated polypeptide for treating, preventing and/or prognosing
 disorders related to the musculoskeletal system including
 musculoskeletal cancers and also for testing and detection e.g.
 diagnosis -
 Example 2; SEQ ID NO 3757; 781bp + Sequence Listing: English.
 The invention relates to novel genes (AAL34669-AAL37666) and proteins
 (AAB03087-AAB04109) associated with the musculoskeletal system useful
 for preventing, treating or ameliorating medical conditions e.g. by
 protein or gene therapy. The genes are isolated from a range of human
 tissues disclosed in the specification. The nucleic acids, proteins,
 antibodies and (ant)agonists are useful in the diagnosis, treatment
 and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 other cancers of the adrenal gland, bone, bone marrow, breast,
 gastrointestinal tract, liver, lung, or urogenital; (b) immune
 disorders e.g. Addison's disease, allergies, autoimmune hemolytic
 anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 and (f) infectious diseases such as viral, bacterial, fungal and
 parasitic infections.
 Note: The sequence data for this patent did not form part of the
 printed specification, but was obtained in electronic format directly
 from WZPO at ftp.wzpo.int/pub/published_pct_sequences.

Sequence 1553 BP; 294 A; 461 C; 563 G; 235 T; 0 other;
 Query Match 20.9%; Score 104.6; DB 22; Length 1553;
 Best Local Similarity 74.9%; Pred. No. 5.8e-12;
 Matches 131; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
 QY 1 GCCCGAAACCCGGAAGTGAACGCGGCGAGCTGCGAGCTCGGAGAAACAGCGCGCGG 60
 Db 921 GCCCGAAACCCGGAAGTGAACGCGGCGAGCTGCGAGCTCGGAGAAACAGCGCGCGG 862
 QY 61 CTCGCGCGCGCGCGCGCGCGCGCGAGATCATGATGCTGCGCGCGCGCGCGCGCGCG 120
 Db 861 CTCGCGCGCGCGCGCGCGCGCGCGAGTGAAGCTCTGCTGCTGCGCGCGCGCGCG 802
 QY 121 GAGCGAGAAAGCCGATGAGACGCCCGCGCGCGCGCGCGCTCTGAGTCCCGCG 175
 Db 801 TCCACCGCGACCCCTGCTGCGCGCGCGCGCGCGCTGCTGCGCGCGCGCGCGCG 747
 RESULT 6
 ID AAA02484/C
 XX AAA02484 standard; cDNA; 1000 BP.
 AC AAA02484;
 DT 19-MAY-2000 (first entry)
 XX
 XX Human colon cancer cell line polynucleotide sequence SEQ ID NO.2475.
 DE
 XX Human; colon cancer; tumour; diagnosis; gene expression product;
 KW probe; detection; cancerous state; metastasis; identification;
 KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
 KW oestrogen receptor-negative breast cancer; lung cancer; ss.
 KW
 XX Homo sapiens.
 OS
 XX
 PN W09558675-A2.
 XX
 XX 18-NOV-1999.
 PD
 XX
 XX 13-MAY-1999; 99NO-US10602.
 PF
 XX
 XX 14-MAY-1998; 98US-0085426.
 PR 15-MAY-1998; 98US-0085337.
 PR 15-MAY-1998; 98US-0085696.
 PR 21-OCT-1998; 98US-0105234.
 PR 27-OCT-1998; 98US-0105877.
 XX
 PA (CHIR) CHIRON CORP.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Leshkowitz D, Kila D, Garcia V, Jones DM, Staehle-Crain B;
 XX
 DR WPI: 2000-126369/11.
 XX
 XX Polynucleotide library used to determine cancerous states of mammalian
 cells -
 Claim 1; Page 994; 1097pp; English.
 AA000010 to AA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived.
 CC The polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a

XX	19-MAY-2000	(first entry)
DE	Human colon cancer cell line polynucleotide sequence SEQ ID NO:2479.	
XX		
XX	Human: colon cancer; tumour; diagnosis; gene expression product;	
KW	probe; detection; cancerous state; metastasis; identification;	
KM	breast cancer; oestrogen receptor-positive breast cancer; therapy;	
XX	oestrogen receptor-negative breast cancer; lung cancer; ss.	
OS	Homo sapiens.	
PN	MO958675-A2.	
PD	18-NOV-1999.	
PF	13-MAY-1999; 99WO-US10602.	
PR	14-MAY-1998; 98US-0085426.	
PR	15-MAY-1998; 98US-0085537.	
PR	15-MAY-1998; 98US-0085696.	
PR	21-OCT-1998; 98US-0105234.	
PR	27-OCT-1998; 98US-0105877.	
PA	(CHIR) CHIRON CORP.	
PA	(HYSE-) HYSEQ INC.	
PI	Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;	
PI	Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;	
PI	Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;	
PI	Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;	
DR	WPI; 2000-126369/11.	
PT	Polynucleotide library used to determine cancerous states of mammalian	
PT	cells -	
PS	Claim 1; Page 995-996; 1097pp; English.	
XX		
XX	AAA00010 to AAA02716 represent polynucleotides isolated from cDNA	
CC	libraries constructed from human colon cancer cell lines. The present	
CC	invention also describes a method of detecting differentially expressed	
CC	genes correlated with a cancerous state of a mammalian cell, comprising	
CC	detecting at least one differentially expressed gene product in a test	
CC	sample derived from a cell suspected of being cancerous, where detection	
CC	of the differentially expressed gene product is correlated with a	
CC	cancerous state of the cell from which the test sample was derived.	
CC	The polynucleotides sequences can be used in a method for detecting	
CC	differentially expressed genes correlated with a cancerous state of a	
CC	mammalian cell. The polynucleotides can also be used as probes for	
CC	detecting and mapping related genes. They can be used in diagnosis and	
CC	prognosis of diseases and disorders (e.g. identification of	
CC	pre-metastatic or metastatic cancerous states, stages of cancer, or	
CC	responsiveness of cancer to therapy). This is particularly for breast	
CC	cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-	
CC	negative breast cancer, lung cancer, and colon cancer.	
SQ	Sequence 1218 BP; 9 A; 31 C; 494 G; 37 T; 647 other;	
OY	Query Match 12.0%; Score 60; DB 21; Length 1218;	
OY	Best Local Similarity 27.6%; Pred. No. 0.0029;	
OY	Matches 123; Conservative 0; Mismatches 323; Indels 0; Gaps 0;	
DB	972 CNCNCNNNNNNNNCCCCCCCCCCCCCNCNCCCNCNNNNNNNNNNNNNNCCCCCCCCCCCC 114	
OY	115 CACCAAGGAGCGAGAAGACCAGATAGACGCCGGGGGGGGTCTGTGAATCCGC 174	
DB	912 CCCANNNNNNNNCCNNCCNNCCNNCCNNCCNNNNNNNNCCNNNNNNNNCCNNCCNNNNNNCC 853	
OY	175 GCCTGCTGCGGGCGGAGAACCCACCCGCGCCGTCGCGCGCATCTTGACAATGGGGCCG 234	
DB	852 CCCNNNNNNNNCCCCCCCCCNCNNCCNNCCCCCCCCCNCNNCCNNCCNNCCNNCCNNCCNNCCCN 793	

[illegible]

CC plasmid pcwEBNA. EBNA 1 protein is used to stably maintain episomes
 CC containing EBV origin of replication (oriP) and a gene encoding
 CC protein or RNA of interest. Eucaryotic host cells expressing EBNA 1
 CC protein are transfected with these episomes to produce recombinant
 CC cell lines expressing multiple genes of interest. This provides a
 CC rapid and reliable method of stably expressing multiple genes in
 CC transfected cells. The episomes are useful in the transfection of genes
 CC encoding receptors, transporter proteins, ion channels, adhesion
 CC molecules and transcription factors. The episomes carrying desired genes
 CC can also be used to transfect cells in gene therapy, antisense therapy,
 CC for gene amplification, cell immortalisation, etc.

XX Sequence 1925 BP; 486 A; 352 C; 872 G; 215 T; 0 other;

Query Match 12.0%; Score 60; DB 20; Length 1925;

Best Local Similarity 50.6%; Pred. No. 0.003;

Matches 170; Conservative 0; Mismatches 165; Indels 1; Gaps 1;

131 GCGCAGATGAGCGCCCGGCGGCGGCTCGAGTCCCGGCGCTGCTGCCGCG 190
 638 GCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
 191 AGACCCCAACCCCGCTGCGCGCGGATGCTGAGTGGGCGCGCATGACAGGATTA 250
 578 GCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
 251 CCGCAGCATGAGAACCCCGGCGGCGGCGGCTGCTGAGCGCGCGCGCTGCGAG 310
 518 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
 311 GCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
 458 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
 370 CCGCAGCATGAGCGCCCGGCGGCGGCGGCTGCTGAGCGCGCGCGCTGCGAG 429
 398 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339
 430 ACACAGCATGAGCGCGCGGCGGCGGCGGCTGCTGAGCGCGCGCTGCGAGCT 465
 338 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303

RESULT 10

ID AAA50254/c standard; DNA; 1926 BP.

XX AAA50254;

DT 07-NOV-2000 (first entry)

XX Epstein Barr virus nuclear antigen 1 protein (EBNA1) DNA.

KW EBV: nuclear antigen 1; EBNA1; episome; transfection; selection;

XX gene therapy; ds.

OS Epstein-barr virus.

PN WO200047778-A1.

PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03547.

PR 11-FEB-1999; 99US-0249585.

PA (PHAR-) PHARMACOPEDIA INC.

PI Horlick RA, Chelsky D;

DR WPI; 2000-515062/46.

XX P-PSDB; AAY95856.

PT Stably transfecting eukaryotic cells with at least one episome for the
 production of a desired protein in vitro and for gene therapy -
 Disclosure: Fig 2; 53pp; English.

XX The present sequence is that of DNA encoding the Epstein-Barr virus
 CC (EBV) nuclear antigen protein 1 (EBNA1, see AAY95856). EBNA1 is
 CC utilised in a novel method for obtaining a eukaryotic cell that is
 CC stably transfected with at least one episome. This method involves
 CC transfecting a eukaryotic cell with: (1) a first episome comprising
 CC an EBV origin of replication (oriP, see AAA50253), a gene encoding a
 CC first protein whose expression results in cell death and a
 CC selectable marker for eukaryotic cells; and (2) a second episome
 CC comprising an EBV oriP and a gene encoding a second protein, where
 CC expression of the second protein prohibits the occurrence of cell
 CC death resulting from expression of the first protein to produce
 CC doubly transfected cells which also express an antigen that
 CC promotes retention of the episomes by the cells. The doubly
 CC transfected cells are maintained under conditions in which the
 CC first and second proteins and the selectable marker are expressed,
 CC and the selective pressure specified by the marker is maintained.
 CC Under these conditions, only cells containing both episomes live.
 CC Preferably, EBNA1 is expressed from 1 of the episomes, and the
 CC protein of interest from the other episome. Either or both
 CC epitopes may further comprise a nucleic acid sequence encoding a
 CC protein desired to be expressed in the cell (e.g. a therapeutic
 CC protein), a nucleic acid encoding an RNA that is not intended to
 CC be translated (e.g. a therapeutic RNA), or a DNA sequence used as
 CC a tag for the cells. The method is applicable to cell culture or
 CC intact organisms, for gene therapy. It allows the rapid
 CC establishment of eukaryotic cells that stably and reliably express
 CC a gene of interest, using a novel method of selection, and
 CC maintenance of that selection without the need for exogenous
 CC selection factors, such as antibiotics.

XX Sequence 1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other;

Query Match 12.0%; Score 60; DB 21; Length 1926;

Best Local Similarity 50.6%; Pred. No. 0.003;

Matches 170; Conservative 0; Mismatches 165; Indels 1; Gaps 1;

131 GCGCAGATGAGCGCCCGGCGGCGGCTCGAGTCCCGGCGCTGCTGCCGCG 190
 638 GCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
 191 AGACCCCAACCCCGCTGCGCGCGGATGCTGAGTGGGCGCGCATGACAGGATTA 250
 578 GCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
 251 CCGCAGCATGAGAACCCCGGCGGCGGCGGCTGCTGAGCGCGCGCGCTGCGAG 310
 518 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
 311 GCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
 458 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
 370 CCGCAGCATGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
 398 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339
 430 ACACAGCATGAGCGCGCGGCGGCGGCGGCTGCTGAGCGCGCGCTGCGAGCT 465
 338 CCGCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303

RESULT 11

ID AAF82902/c standard; DNA; 1926 BP.

XX AAF82902;

DT 29-JUN-2001 (first entry)

Fri Oct 11 09:30:40 2002

us-09-489-101a-13_copy_1_500.png

Job time : 103 secs

Fri Oct 11 09:30:38 2002

us-09-489-101a-12_copy_1_500.rni

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 18.733 Seconds
(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101a-12_COPY_1_500

Sequence: 1 aagcttggtgcattatc.....ccctgcgcggagcagaag 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCBUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/PCBUseq1.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	31.8	6.4	US-08-287-442-3	Sequence 3, Appl1
2	31.8	6.4	US-08-459-701-3	Sequence 3, Appl1
3	31.8	6.4	US-08-460-298-3	Sequence 3, Appl1
4	31.8	6.4	US-08-459-174-3	Sequence 3, Appl1
5	31.8	6.4	PCT-US93-06300A-3	Sequence 3, Appl1
6	31.8	6.4	US-08-761-258-1	Sequence 1, Appl1
7	31.8	6.4	US-08-977-306-1	Sequence 1, Appl1
8	30	6.0	US-08-343-443B-106	Sequence 106, App
9	30	6.0	US-08-343-443B-1	Sequence 1, Appl1
10	29.6	5.9	US-09-141-000-2	Sequence 2, Appl1
11	29.6	5.9	US-09-103-840A-2	Sequence 2, Appl1
12	29	5.8	US-08-050-132A-8	Sequence 8, Appl1
13	29	5.8	US-08-750-232A-8	Sequence 8, Appl1
14	29	5.8	US-08-815-652A-8	Sequence 8, Appl1
15	29	5.8	US-08-254-353A-8	Sequence 8, Appl1
16	29	5.8	PCT-US92-05374A-8	Sequence 8, Appl1
17	29	5.8	PCT-US95-07084-8	Sequence 8, Appl1
18	29	5.8	US-08-815-718-1	Sequence 1, Appl1
19	29	5.8	US-08-884-072-4	Sequence 4, Appl1
20	29	5.8	US-09-213-168-4	Sequence 4, Appl1
21	29	5.8	US-08-368-043-1	Sequence 1, Appl1
22	28.8	5.8	US-08-012-543-1	Sequence 1, Appl1
23	28.8	5.8	PCT-US93-07645A-1	Sequence 1, Appl1
24	28.8	5.8	PCT-US93-07645-1	Sequence 1, Appl1
25	28.8	5.8	US-08-371-121-15	Sequence 15, Appl1
26	28.8	5.8	US-09-127-424-1	Sequence 1, Appl1
27	28.8	5.8	US-09-495-082-1	Sequence 1, Appl1

28	28.8	5.8	3705	5	PCT-US96-03940-7	Sequence 7, Appl1
29	28.8	5.8	5648	5	PCT-US96-03940-1	Sequence 1, Appl1
30	28.6	5.7	838	4	US-09-007-097-1	Sequence 1, Appl1
31	28.6	5.7	838	5	PCT-US94-10457-1	Sequence 1, Appl1
32	28.6	5.7	838	5	PCT-US96-03965-7	Sequence 7, Appl1
33	28.6	5.7	946	2	US-08-816-605-1	Sequence 1, Appl1
34	28.6	5.7	1415	4	US-08-236-918A-7	Sequence 7, Appl1
35	28.6	5.7	1415	4	US-09-150-864A-7	Sequence 142, App
36	28.4	5.7	662	4	US-08-936-165A-142	Sequence 4, Appl1
37	28.4	5.7	1035	1	US-07-975-526-4	Sequence 5, Appl1
38	28.4	5.7	5438	4	US-08-456-200B-5	Sequence 11, Appl1
39	28.4	5.7	6232	4	US-08-456-200B-11	Sequence 8, Appl1
40	28.4	5.7	7175	2	US-08-455-543A-8	Sequence 8, Appl1
41	28.4	5.7	7175	2	US-08-193-078B-8	Sequence 8, Appl1
42	28.4	5.7	7175	2	US-08-223-305C-8	Sequence 8, Appl1
43	28.4	5.7	7175	2	US-08-149-097D-8	Sequence 8, Appl1
44	28.4	5.7	7175	3	US-08-949-386-8	Sequence 8, Appl1
45	28.4	5.7	7175	3	US-08-450-562-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-287-442-3
; Sequence 3, Application US/08287442
; Patent No. 5670350
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas D.
; APPLICANT: Lam, Stephen T.
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Stein, Jeffrey I.
; APPLICANT: Howell, Charles R.
; APPLICANT: Becker, J. Ole
; TITLE OF INVENTION: Gene Activating Element
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,442
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,636
; FILING DATE: 01-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,284
; FILING DATE: 02-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/570,184
; FILING DATE: 08-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5559 base pairs
; type: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudomonas fluorescens
STRAIN: CGA267356
INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
IMMEDIATE SOURCE: fragment
CLONE: pCIB137
FEATURE:
NAME/KEY: misc-feature
LOCATION: 210..1688
OTHER INFORMATION: /note= "ORF 1, transcribed left to
right"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1906..3633
OTHER INFORMATION: /note= "ORF 2, transcribed left to
right"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 4616..4691
OTHER INFORMATION: /note= "glyw, transcribed right to
left"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 4731..5318
OTHER INFORMATION: /note= "ORF 3, transcribed right to
left"
US-08-287-442-3

Query Match
Best Local Similarity 50.3%; Score 31.8; DB 1; Length 5559;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 109 GATTGGTGTGATTTGATGTCGACGAGTGAAGTAAATGCTTATCTGCTGACGCTG 168
Db 1198 GGTTCCTGCGCTATGTGCGGGGGAATCAACCCGACCTGCTGACCTGATCCAGCCCG 1257
Qy 169 TGTGTGAGATGGCTGTGTCTGCCACCTCTCGAGTGCATTTTGCATGTGTAACAGGT 228
Db 1358 AGCTGGCCCTTGAACTGGCGACCAAGCTGTTCAGGTGCAACAGTCCGATGCGGTGA 1317
Qy 229 CTCCTCTGTGGGCAACCAACAAAGAGAACTTGC 263
Db 1318 CCGCCGCGGCGGATCGATCGACGGGAAAGAAC 1352

RESULT 2
US-08-459-701-3
Sequence 3, Application US/08459701
Patent No. 5686282
GENERAL INFORMATION:
APPLICANT: Gaffney, Thomas D.
APPLICANT: Lam, Stephen T.
APPLICANT: Lyon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Stein, Jeffrey I.
APPLICANT: Howell, Charles R.
APPLICANT: Becker, J Ole
TITLE OF INVENTION: Gene Activating Element
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,701
FILING DATE: 02-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/287,442
FILING DATE: 08-AUG-1994
APPLICATION NUMBER: US 08/087,636
FILING DATE: 01-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,284
FILING DATE: 02-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/570,184
FILING DATE: 08-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEO ID NO: 3;
SEQUENCE CHARACTERISTICS:
LENGTH: 5559 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudomonas fluorescens
STRAIN: CGA267356
INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
IMMEDIATE SOURCE: fragment
CLONE: pCIB137
FEATURE:
NAME/KEY: misc-feature
LOCATION: 210..1688
OTHER INFORMATION: /note= "ORF 1, transcribed left to
right"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1906..3633
OTHER INFORMATION: /note= "ORF 2, transcribed left to
right"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 4616..4691
OTHER INFORMATION: /note= "glyw, transcribed right to
left"
FEATURE:
NAME/KEY: misc-feature
LOCATION: 4731..5318
OTHER INFORMATION: /note= "ORF 3, transcribed right to
left"
US-08-459-701-3

Query Match
Best Local Similarity 50.3%; Score 31.8; DB 1; Length 5559;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 109 GATTGGTGTGATTTGATGTCGACGAGTGAAGTAAATGCTTATCTGCTGACGCTG 168
Db 1198 GGTTCCTGCGCTATGTGCGGGGGAATCAACCCGACCTGCTGACCTGATCCAGCCCG 1257
Qy 169 TGTGTGAGATGGCTGTGTCTGCCACCTCTCGAGTGCATTTTGCATGTGTAACAGGT 228

Db 1258 AGCTGCGCCCTTGGACTGCGCCACGCGCTGTCACGACGTCGCTGCGGTGA 1317
OY 229 CTCCCTCTGTGGGCGACACACAAAGAGAGTTGC 263
Db 1318 CCGCGCCGCGGCTGCGATCGACGCGGGAAGAAGC 1352

RESULT 3
US-08-460-298-3
Sequence 3, Application US/08460298
Patent No. 5686283
GENERAL INFORMATION:
APPLICANT: Gaffney, Thomas D.
APPLICANT: Lam, Stephen T.
APPLICANT: Liqon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Stein, Jeffrey I.
APPLICANT: Howell, Charles R.
APPLICANT: Becker, J. Ole
TITLE OF INVENTION: Gene Activating Element
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/287,442
FILING DATE: 08-AUG-1994
APPLICATION NUMBER: US 08/087,636
FILING DATE: 01-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,284
FILING DATE: 02-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/570,184
FILING DATE: 08-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5559 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
AMBI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: pseudomonas fluorescens
STRAIN: CGA267356
INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
INDIVIDUAL ISOLATE: fragment
IMMEDIATE SOURCE:
CLONE: pCIB137
FEATURE:

NAME/KEY: misc.feature
LOCATION: 210..1688
OTHER INFORMATION: /note="ORF 1, transcribed left to
OTHER INFORMATION: right"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1906..3633
OTHER INFORMATION: /note="ORF 2, transcribed left to
OTHER INFORMATION: right"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 4616..4691
OTHER INFORMATION: /note="g1yw, transcribed right to
OTHER INFORMATION: left"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 4731..5318
OTHER INFORMATION: /note="ORF 3, transcribed right to
OTHER INFORMATION: left"
US-08-460-298-3

Query Match 6.4%; Score 31.8; DB 1; Length 5559;
Best local similarity 50.3%; Pred. No. 1.3;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

OY 109 GATGTGTGATTTGATGCTGACGGAGAGATGAGATGAGCTTATGCTGACGCTG 168
Db 1198 GTTCTGCTGCTATGTCGCGGGGGAATCACCACCTGCTGACCTGATCCACCCG 1257
OY 169 TGCTGAGATGCGCCGCTGCTGCGCCACCTCTCTGAGTACATTTTGCATGTACAGG 228
Db 1258 AGCTGCGCCCTTGAACCTGCGACGACGCTGTTCACAGTCCACACTCGGCTGCGGTGA 1317
OY 229 CTCCCTCTGTGGGCGACACAAAGAGAGTTGC 263
Db 1318 CCGCGCCGCGGCTGCGATCGACGCGGGAAGAAGC 1352

RESULT 4
US-08-459-174-3
Sequence 3, Application US/08459174
Patent No. 5710031
GENERAL INFORMATION:
APPLICANT: Gaffney, Thomas D.
APPLICANT: Lam, Stephen T.
APPLICANT: Liqon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Stein, Jeffrey I.
APPLICANT: Howell, Charles R.
APPLICANT: Becker, J. Ole
TITLE OF INVENTION: Gene Activating Element
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,174
FILING DATE: 02-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/287,442
FILING DATE: 08-AUG-1994
APPLICATION NUMBER: US 08/087,636
FILING DATE: 01-JUL-1993

Fri Oct 11 09:30:38 2002

us-09-489-101a-12_copy_1_500.rn1

Page 4

```

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: US 07/908,284
3 FILING DATE: 02-JUL-1992
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: US 07/570,184
6 FILING DATE: 08-AUG-1990
7 ATTORNEY/AGENT INFORMATION:
8 NAME: Elmer, James Scott
9 REGISTRATION NUMBER: 36,129
10 REFERENCE/DOCKET NUMBER: 36,129
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: 919-541-8614
13 TELEFAX: 919-541-8689
14 INFORMATION FOR SEQ ID NO: 3:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 5559 base pairs
17 TYPE: nucleic acid
18 STRANDEDNESS: single
19 TOPOLOGY: linear
20 MOLECULE TYPE: DNA (genomic)
21 HYPOTHEetical: NO
22 ANTI-SENSE: NO
23 ORIGINAL SOURCE:
24 ORGANISM: Pseudomonas fluorescens
25 STRAIN: CGA267356
26 INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
27 INDIVIDUAL SOURCE: fragment
28 IMMEDIATE SOURCE:
29 CLONE: pCIB137
30 FEATURE:
31 NAME/KEY: misc_feature
32 LOCATION: 210..1688
33 OTHER INFORMATION: /note="ORF 1, transcribed left to
34 OTHER INFORMATION: right"
35 FEATURE:
36 NAME/KEY: misc_feature
37 LOCATION: 1906..3633
38 OTHER INFORMATION: /note="ORF 2, transcribed left to
39 OTHER INFORMATION: right"
40 FEATURE:
41 NAME/KEY: misc_feature
42 LOCATION: 4616..4691
43 OTHER INFORMATION: /note="g1yW, transcribed right to
44 OTHER INFORMATION: left"
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: 4731..5318
48 OTHER INFORMATION: /note="ORF 3, transcribed right to
49 OTHER INFORMATION: left"
50 US-08-459-174-3
51
52 Query Match 6.4%; Score 31.8; DB 1; Length 5559;
53 Best Local Similarity 50.3%; Pred. No. 1.3;
54 Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
55
56 QY 109 GATTGATGTGATTTGAGTCTGACGAGAGTGAAGTATGCGCTTATCCGTCTGCAGGCTG 168
57 Db 1198 GGTTCCTCTCGCTATGTGTGGCGGGGAAATACCCGCCACCTGCTGACCCGCG 1257
58
59 QY 169 TCGTACGATGCGCTGCTGTGCGACCCCTCTCGAGTGAACATTTTGGATGTGAACAGAGGT 228
60 Db 1258 AGCTGCGCTTGACACTCGCACACGACGCTGTTCACGATGACAGACAGTCCGCTGTGGTGGGA 1317
61
62 QY 229 CTCCTCTGGGGCACAACAACAAGAGAGATGCG 263
63 Db 1318 CCGCCGCGGGTGGCATCGACGCGGAGAAAGAAC 1352
64
65 RESULT 5
66 PCT-US93-06300A-3
67 Sequence 3, Application PC/TUS9306300A
68 GENERAL INFORMATION:
69 APPLICANT: CIBA-GEIGY AG

```

```

1  APPLICANT: Rlybeckstrasse 141
2  APPLICANT: 4002 Basle
3  APPLICANT: Switzerland
4  APPLICANT: 125 Tradescant Road
5  APPLICANT: Chapel Hill, NC 27514
6  APPLICANT: USA
7  APPLICANT: 8900 Jeanne Court
8  APPLICANT: Raleigh, NC 27613
9  APPLICANT: USA
10 APPLICANT: Hill, Dwight Steven
11 APPLICANT: 311 Melanite Lane
12 APPLICANT: Cary, NC 27511
13 APPLICANT: USA
14 APPLICANT: Stein, Jeffrey I.
15 APPLICANT: 3725 Surry Trail
16 APPLICANT: Hillsborough, NC 27278
17 APPLICANT: USA
18 APPLICANT: Howell, Charles R.
19 APPLICANT: 805 Avondale
20 APPLICANT: Bryan, TX 77802
21 APPLICANT: USA
22 APPLICANT: Becker, J. Ole
23 APPLICANT: 6164 Osevego
24 APPLICANT: Riverside, CA 92506
25 APPLICANT: USA
26 APPLICANT: Ligon, James M.
27 APPLICANT: 120 Marquette Drive
28 APPLICANT: Cary, NC 27513
29 APPLICANT: USA
30 APPLICANT:
31 TITLE OF INVENTION: Gene Activating Element
32 NUMBER OF SEQUENCES: 7
33 CORRESPONDENCE ADDRESS:
34 ADDRESSEE: CIBA-GEIGY Corporation
35 STREET: 7 Skyline Drive
36 CITY: Hawthorne
37 STATE: New York
38 COUNTRY: USA
39 ZIP: 10632
40 COMPUTER READABLE FORM:
41 MEDIUM TYPE: floppy disk
42 OPERATING SYSTEM: IBM PC compatible
43 SOFTWARE: Patent In Release #1.0, Version #1.25
44 CURRENT APPLICATION DATA:
45 APPLICATION NUMBER: PCT/US93/06300A
46 FILING DATE: 02-JUL-1993
47 CLASSIFICATION:
48 PRIOR APPLICATION DATA:
49 APPLICATION NUMBER: US 07/908,284
50 FILING DATE: 02-JUL-1992
51 ATTORNEY/AGENT INFORMATION:
52 NAME: Sprull, W. Murray
53 REGISTRATION NUMBER: 32,943
54 REFERENCE/DOCKET NUMBER: S-18210/A/CGC1506/PC
55 TELECOMMUNICATION INFORMATION:
56 TELEPHONE: (919)541-8615
57 TELEFAX: (919)541-8689
58 INFORMATION FOR SEQ ID NO: 3:
59 SEQUENCE CHARACTERISTICS:
60 LENGTH: 5559 base pairs
61 TYPE: nucleic acid
62 STRANDEDNESS: single
63 TOPOLOGY: linear
64 MOLECULE TYPE: DNA (genomic)
65 HYPOTHEetical: NO
66 ANTI-SENSE: NO
67 ORIGINAL SOURCE:
68 ORGANISM: Pseudomonas fluorescens
69 STRAIN: Cgk267356
70 INDIVIDUAL ISOLATE: 5.6 kb EcoRI-HindIII restriction
71 INDIVIDUAL ISOLATE: Fragment
72 IMMEDIATE SOURCE:
73 CLONE: pcIB137

```

```

1 TELEFAX: (919) 541-8689
2 INFORMATION FOR SEQ ID NO: 1:
3 SEQUENCE CHARACTERISTICS:
4     LENGTH: 10763 base pairs
5     TYPE: nucleic acid
6     STRANDEDNESS: single
7     TOPOLOGY: linear
8     MOLECULE TYPE: DNA (genomic)
9     HYPOTHETICAL: NO
10    ANTI-SENSE: NO
11    ORIGINAL SOURCE:
12     ORGANISM: Pseudomonas fluorescens
13     STRAIN: CGA267356 (aka MOC134 and aka BL915)
14 IMMEDIATE SOURCE:
15     CLONE: plasmid pE11
16 FEATURE:
17     NAME/KEY: misc_feature
18     LOCATION: 210..1688
19     OTHER INFORMATION: /product= "methyltransferase"
20     OTHER INFORMATION: /note= "Coding sequence for methyltransferase has homology to
21     OTHER INFORMATION: the cscA and ftrp genes from E. coli and Myxococcus xanthus"
22     OTHER INFORMATION: respectively."
23 FEATURE:
24     NAME/KEY: misc_feature
25     LOCATION: 1906..3633
26     OTHER INFORMATION: /product= "sensor kinase"
27     OTHER INFORMATION: /note= "Coding sequence for sensor kinase has homology to
28     OTHER INFORMATION: cscC, ftrp, and bvgS genes of E. coli, M. xanthus, and
29     OTHER INFORMATION: Bordetella pertussis, respectively."
30 FEATURE:
31     NAME/KEY: misc_RNA
32     LOCATION: complement (4616..4691)
33     OTHER INFORMATION: /product= "tRNA"
34     OTHER INFORMATION: /note= "(complementary DNA strand) Homology to glyW from E. coli"
35     OTHER INFORMATION: Coli."
36 FEATURE:
37     NAME/KEY: misc_feature
38     LOCATION: complement (4731..5318)
39     OTHER INFORMATION: /product= "response
40     OTHER INFORMATION: "CDP-diacylglycerol-3-phosphate-3-phosph
41     OTHER INFORMATION: atidylitran."
42     OTHER INFORMATION: /note= "Coding sequence for
43     OTHER INFORMATION: CDP-diacylglycerol-3-phosphate-3-phosphatidyltran
44     OTHER INFORMATION: se has homology to pgsA."
45 FEATURE:
46     NAME/KEY: misc_feature
47     LOCATION: complement (5574..7397)
48     OTHER INFORMATION: /product= "UVR exonuclease subunit
49     OTHER INFORMATION: C"
50     OTHER INFORMATION: /note= "Coding sequence for UVR exonuclease subunit C has
51     OTHER INFORMATION: homology to uvrC."
52 FEATURE:
53     NAME/KEY: misc_feature
54     LOCATION: complement (7400..8041)
55     OTHER INFORMATION: /function= "response
56     OTHER INFORMATION: regulator/transcription activator"
57     OTHER INFORMATION: /product= "gacA (aka gafA)"
58     OTHER INFORMATION: /note= "Coding sequence for gacA (aka gafA) has homology to
59     OTHER INFORMATION: uvvY and gacA genes of E. coli and Ps. fluorescens,
60     OTHER INFORMATION: respectively."
61
62 US-08-761-256-1
63
64 Query Match
65     Best Local Similarity 6.4%: Score 31.8; DB 1; Length 10763;
66     Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
67
68 Oy 109 GATTCGTGTGATTTCATGTCGACGAGGAGTAAATGSCCTTATCCGTCGCGAGGCTG 168
69 Db 1198 GGGTCTCTGCGATATGTCGCGGGGGAATACCCGACACCTGCTGACCTGATCAGCCGC 1257
70
71 169 TCGTAGAGATGGCTGTGTCGCCACCCCTCTCGAGTAGCAATTTGCAATGCTGTAACAGGT 228
72 1258 AGCTGGCGCTTGAACTCGACGACGACGCTTTCACAGGTGCAACAGCTCCGGTGTGGGGGA 1317

```


QY 229 CTCCTCTGGGACACACACAGAGAGTTGC 263
DB 1318 CCGCCGGGGGTGCATCGACCGGAGAGAGC 1352

RESULT 7
US-08-977-306-1
Sequence 1, Application US/08977306
Patent No. 595348
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Gaffney, Thomas D.
APPLICANT: Torkewitz, Nancy
APPLICANT: Stafford, Jill M.
TITLE OF INVENTION: Genetically Modified Pseudomonas Strains
TITLE OF INVENTION: With Enhanced Biocontrol Activity
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 595348artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/977,306
FILING DATE:
CLASSIFICATION: A35
AUTHOR/AGENT INFORMATION:
NAME: Meigs, T. Timothy
REGISTRATION NUMBER: 36,241
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10763 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudomonas fluorescens
STRAIN: CGA267356 (aka MOC6134 and aka BL915)
IMMEDIATE SOURCE:
CLONE: Plasmid pE11
FEATURE:
NAME/KEY: misc-feature
LOCATION: 210..1688
OTHER INFORMATION: /product="methyltransferase"
OTHER INFORMATION: /note="Coding sequence for methyltransferase has homology to
OTHER INFORMATION: the cheR and frz genes from E. coli and Myxococcus xanthus,
OTHER INFORMATION: respectively."
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1906..3633
OTHER INFORMATION: /product="sensor kinase"
OTHER INFORMATION: /note="Coding sequence for sensor kinase has homology to the
OTHER INFORMATION: rscG, frzE, and bzg genes of E. coli, M. xanthus, and
OTHER INFORMATION: Bordetella pertussis, respectively."
NAME/KEY: misc-RNA
LOCATION: complement (4616..4691)
OTHER INFORMATION: /product="tRNA"

OTHER INFORMATION: /note="(complementary DNA strand) Homology to glyw from E
OTHER INFORMATION: coli."
FEATURE:
NAME/KEY: misc-feature
LOCATION: complement (4731..5318)
OTHER INFORMATION: /product="CDP-diacylglycerol-3-phosphate-3-phosph
OTHER INFORMATION: atidylyltrans."
OTHER INFORMATION: /note="Coding sequence for
OTHER INFORMATION: CDP-diacylglycerol-3-phosphate-3-phosphatidylyltran
OTHER INFORMATION: se has homology to pgsa."
FEATURE:
NAME/KEY: misc-feature
LOCATION: complement (5574..7397)
OTHER INFORMATION: /product="uvr exonuclease subunit
OTHER INFORMATION: C"
OTHER INFORMATION: /note="Coding sequence for uvr exonuclease subunit C has
OTHER INFORMATION: homology to uvrC."
FEATURE:
NAME/KEY: misc-feature
LOCATION: complement (7400..8041)
OTHER INFORMATION: /function="response
OTHER INFORMATION: regulator/transcription activator"
OTHER INFORMATION: /product="gacA (aka gafa)"
OTHER INFORMATION: /note="Coding sequence for gacA (aka gafa) has homology to
OTHER INFORMATION: uvrI and gacA genes of E. coli and Ps. fluorescens,
OTHER INFORMATION: respectively."

US-08-977-306-1
Query Match 6.4%; Score 31.8; DB 2; Length 10763;
Best Local Similarity 50.3%; Pred. No. 1.9; 77; Indels 0; Gaps 0;
Matches 78; Conservative 0; Mismatches 0

QY 109 GATTGTGTCATTTGATGCTGACGGAGTGAAGTAATGACCTTATCTGTCACAGGCTG 168
DB 1198 GTTTCCTGCGCCCTATGTCGGGAAATCACCCGACCTGACCTGATCCAGCCCG 1257
QY 169 TGTCTAGGATGCGCTGCTGTCGCCACCTCTCTCGAGTAGCATTTTTCAGTGTACAGGGT 228
DB 1258 AGCTGCGCCCTTGAACCTGCGCACACAGCTGTTCCAGGTGCAACAGTGGTGTGGCGTGA 1317
QY 229 CTCCTCTGGGACACACACAGAGAGTTGC 263
DB 1318 CCGCCGGGGGTGCATCGACCGGAGAGAGC 1352

RESULT 8
US-08-343-443B-106/c
Sequence 106, Application US/08343443B
Patent No. 5968734
GENERAL INFORMATION:
APPLICANT: Aurias, Alain
APPLICANT: Delattre, Olivier
APPLICANT: Desmaze, Chantal
APPLICANT: Melot, Thomas
APPLICANT: Peter, Martine
APPLICANT: Ploquaestel, Beatrice
APPLICANT: Zucman, Jacques
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TITLE OF INVENTION: TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
TITLE OF INVENTION: TRANSLOCATIONS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443.443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 954 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..885
US-08-343-443B-106

Query Match
Best Local Similarity 67.7%; Score 30; DB 2; Length 954;
Matches 42; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
DB 154 GTCGGGAAATCTCCAGACTCTGCGCATTAACACCCACTGCTGCTGTC 95
QY 475 GG 476
DB 94 GG 93

RESULT 9
US-08-343-443B-1/c
Sequence 1, Application US/08343443B
Patent No. 5968734
GENERAL INFORMATION:
APPLICANT: Aulias, Alain
APPLICANT: Delatire, Olivier
APPLICANT: Desmazes, Chantal
APPLICANT: Melot, Thomas
APPLICANT: Peter, Martine
APPLICANT: Ploougestel, Beatrice
APPLICANT: Thomas, Gilles
APPLICANT: Zucman, Jessica
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443.443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2371 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1992
US-08-343-443B-1

Query Match
Best Local Similarity 67.7%; Score 30; DB 2; Length 2371;
Matches 42; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
DB 886 GTCGGGAAATCTCCAGACTCTGCGCATTAACACCCACTGCTGCTGTC 827
QY 475 GG 476
DB 826 GG 825

RESULT 10
US-09-141-000-2/c
Sequence 2, Application US/09141000
Patent No. 6054295
GENERAL INFORMATION:
APPLICANT: Chen, Pang
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
TITLE OF INVENTION: RECEPTOR PROTEINS
FILE REFERENCE: 199991
CURRENT APPLICATION NUMBER: US/09/141.000
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Pasted for Windows Version 3.0
SEQ ID NO 2
LENGTH: 500
TYPE: PRT
ORGANISM: Human
US-09-141-000-2
Query Match
Best Local Similarity 5.9%; Score 29.6; DB 3; Length 500;
Matches 38; Conservative 132; Mismatches 284; Indels 0; Gaps 0;
QY 44 TTAGGGAACATTTGTCAGCAAAAGTAAATAGCAAACTAGCCTTTATTTTA 103
DB 456 TTD..S.B...TW...SHYTWIT..K...MDK.BMC..MBSR.D.BMTTA.Y... 397
QY 104 AATTGATGCTGCTGATTCGATTCGACGCGGAGTGAGTATGCGCTTCTGCTGCA 163

Db 396 . . . MCRYW . . . DY . S . . RH . T . D . H . M . BT . H . KRSHSN . T . TM . AB . . . M . 337
 Qy 164 GCGTGTGCTGAGGATGCGCTGTGCTGGCACCCTCCGACGAGTAGCATTTTGATGTGTATAC 222
 Db 336 BM.MMYRYB . . . TYR . . . CT . YSD . HK . RH . TFB . MHHR . SYNB . C KWS . 277
 Qy 224 AAGGTCCTCCCTGTGGGGGACACACACAGAGAGATTTGTTAAGACAGAGAGCGTC 268
 Db 276 . . . SK . HC . S . SS . C . DMTWC . BB . YH . HC . AA . TM . HC . KC . . . KTR . MH . TB . 217
 Qy 284 GGAATGATCATCTCCATCTGAGAGACGCTGGGTATACCCAGATGCTGAGAGAGAGTGT 343
 Db 216 . RSB . MA . MMT . S . . . S . R . SS . SH . YMYR . YCTBYH . YBC . M . KCBM . GMR . YGT 157
 Qy 344 TGGCAGGTCCTCCACAGAGCTTCGACGCTGACCTTGGGGGTGAGACATCTGTCTGTTC 403
 Db 156 . GSNYYMA . G . MAT . GSR . NC . . . AYV . . TMGT . GSBGBCDKST . KGB . G . TM . T 97
 Qy 404 TCGCGATTAAGCGCGCTGAAAGCAGCCACATCTCTCCCAATATACACAGCGATTGG 463
 Db 96 N . KR . GM . TSH . K . CSTGH . NMTE . ACC . CTGT . K . . SH . CNTDAC . T . C . CTSHTSSC 37
 Qy 464 GGTTCCTCATTGGCGCACCTTCGCGGAGGCCAAG 497
 Db 36 S . SDDS . TH . CSS . SS . . . AM . . SCSSC . DYHNS 3

RESULT 11
US-09-103-840A-2/C
Sequence 2, Application US/09103840A

```

1 GENERAL INFORMATION:
2 APPLICANT: FLEISCHMAN, Robert D.
3 APPLICANT: WHITE, Owen R.
4 APPLICANT: FRASER, Claele M.
5 APPLICANT: VENER, John C.
6 TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
7 TITLE OF INVENTION: TUBERCULOSIS
8 TITLE OF INVENTION: 1366-4407, DO
9 CURRENT PUBLICATION NUMBER: 55/09/103,840A
10 CURRENT FILING DATE: 1998-06-24
11 NUMBER OF SEQ ID NOS: 2
12 SOFTWARE: PatentIn Ver. 2.1
13 SEQ ID NO. 2
14 LENGTH: 4403765
15 TYPE: DNA
16 ORGANISM: Mycobacterium tuberculosis
17 FEATURE:
18 OTHER INFORMATION: CDC 1551
19 OTHER INFORMATION: "n" bases at various positions throughout the sequence
20 OTHER INFORMATION: represent a, t, c or g
21 IS-09-103-840A-2

```

Query Match	5.9%	Score 29.6;	DB 4;	Length 4403765;
Best Local Similarity	52.2%;	Pred. No. 54;		
Matches 70; Conservative	0;	Mismatches 64;	Indels 0;	Gaps 0;

QY 315 GCTTACTCCAAATGGCTGAGAGAGGTGATAGGCAATGCTCCCAAGAGCGTCAGAGCGCA 374
Db 514674 GTTGACAGCCTATGCGGGCGTTGGTGCAAATGGCCAAATGGGGGTGGGGTTGGCCGGGTGGA 5146155
QY 375 CTTGGGGGGGAGCACTCTCGTCGTCCTTCCGCTGATTAACGGCGCTGAAACCGACCCAC 434
Db 514614 CCGGCGTATGTCGCTTGGGTGCTCACCGGAAACTCTATTCTTGGCGGCGACAGGCGCT 5145555
QY 435 TGCTGCCCAAAATC 448
Db 514554 GCGTGGCGAAAGTC 514541

RESULT 12
US-08-050-132A-8/C
; Sequence 8, Application US/08050132A

Patent No. 5661007

 GENERAL INFORMATION:
 APPLICANT: Mooney, John M.
 APPLICANT: Celestey, Anthony
 TITLE OF INVENTION: BMP-9 COMPOSITIONS
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genetics Institute, Inc.
 STREET: Legal Affairs - 87 Cambridgepark Drive
 CITY: Cambridge
 STATE: MA
 COUNTRY: US
 ZIP: 02140

 COMPUTER RELEVABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/050,132A
 FILING DATE:

 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kapinos, Ellen J.
 REGISTRATION NUMBER: 32,245
 REFERENCE/DOCKET NUMBER: GI 5186A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 876-1170
 TELEFAX: (617) 876-5851

 INFORMATION FOR SEQ. ID NO.: 8:

```

1 SEQUENCE CHARACTERISTICS:
2     LENGTH: 470 base pairs
3     TYPE: nucleic acid
4     STRANDEDNESS: double
5     TOPOLOGY: linear
6     MOLECULE TYPE: DNA (genomic)
7     HYPOTHETICAL: NO
8     FRAGMENT TYPE: C-terminal
9     ORIGINAL SOURCE:
10        ORGANISM: Homo sapiens
11        CELL LINE: Wi38 (genomic DNA)
12     IMMEDIATE SOURCE:
13        LIBRARY: human genomic library
14        CLONE: lambda 111-1
15     POSITION IN GENOME:
16        UNITS: bp
17     FEATURE:
18        NAME/KEY: exon
19        LOCATION: 1..470
20     FEATURE:
21        NAME/KEY: CDS
22        LOCATION: 1..456
23     FEATURE:
24        NAME/KEY: mat_peptide
25        LOCATION: 124..453
26     FEATURE:
27        NAME/KEY: mRNA
28        LOCATION: 1..470
29 JS-08-050-132A-8

```

Query Match	5.8%;	Score 29;	DB 1;	Length 470;
Best Local Similarity	52.0%;	Pred. No. 2.9;		
Matches 65; Conservative	0;	Mismatches 60;	Indels 0;	Gaps 0;

QY	333	AGAGGGCTGTATGAGCGAGTCTCCAGACACTGTGACGTCACTATTTGGGGGAGGACATTC	391
Db	347	ACACGAGGAGGCTTGTGCCACCTTTGTGGGAACTTGTAGATCTCACACAGGTCTGTACAGATA	288
QY	392	TGCGTCTTGTCTGCGTGTATTAAGGCGCGTGAAGACGAGCCAACTGTCTGCCCAAAATCAC	451
Db	287	GGGTCTTGTCTGGGCGTCACTGTGTACGCGCAAGGGGGAAGAGCAGCGCGCTTTACACTCG	228
QY	452	CAGCC	456

LOCATION: 1..456
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 124..453
FEATURE:
NAME/KEY: mRNA
LOCATION: 1..470
US-08-815-652B-8

Query Match
Best local Similarity 52.0%; Score 29; DB 3; Length 470;
Matches 65; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 332 AGAGAGTCTATGCGCAGTCTCTCCAGAGCTCTGCGAGCTCTGGGGGTGACAGTC 391
DB 347 ACACAGAGGCGCTTGCCACCTTTGTGGGGAATGAGATGACACAGGCTCTGACGATA 288
OY 392 TCGTGTCTGCTGCGGTGATTAAGGCGCGTGAAGCCAGCCACTGCTGCCAAATCACC 451
DB 287 GCGGTCTCTGCGGCTCACATGTCAGCCAAAGGGGAAGAGCAGCGCCCTTACACTCG 228
OY 452 CAGCC 456
DB 227 TAGGC 223

RESULT 15
US-08-254-353A-8/c
Sequence 8, Application US/08254353A
Patent No. 6287816
GENERAL INFORMATION:
APPLICANT: Rosen, Vicki A.
APPLICANT: Mooney, John M.
APPLICANT: Celeste, Anthony J.
APPLICANT: Song, Jeffrey
APPLICANT: Thies, Scott
TITLE OF INVENTION: BMP-9 COMPOSITIONS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: Legal Affairs - 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: US
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,353A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,245
REFERENCE/DOCKET NUMBER: GI 51868
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 470 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEetical: NO
FRAGMENT TYPE: C-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL LINE: W138 (genomic DNA)

IMMEDIATE SOURCE:
LIBRARY: human genomic library
CLONE: Lambda 111-1
POSITION IN GENOME:
UNITS: bp
FEATURE:
NAME/KEY: exon
LOCATION: 1..470
FEATURE:
NAME/KEY: CDS
LOCATION: 1..456
NAME/KEY: mat_peptide
LOCATION: 124..453
FEATURE:
NAME/KEY: mRNA
LOCATION: 1..470
US-08-254-353A-8

Query Match
Best local Similarity 52.0%; Score 29; DB 4; Length 470;
Matches 65; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 332 AGAGAGTCTATGCGCAGTCTCTCCAGAGCTCTGCGAGCTCTGGGGGTGACAGTC 391
DB 347 ACACAGAGGCGCTTGCCACCTTTGTGGGGAATGAGATGACACAGGCTCTGACGATA 288
OY 392 TCGTGTCTGCTGCGGTGATTAAGGCGCGTGAAGCCAGCCACTGCTGCCAAATCACC 451
DB 287 GCGGTCTCTGCGGCTCACATGTCAGCCAAAGGGGAAGAGCAGCGCCCTTACACTCG 228
OY 452 CAGCC 456
DB 227 TAGGC 223

Search completed: October 10, 2002, 19:26:08
Job time : 2506.73 secs

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="264314"
/clone_lib="CTBT-E1"
/sex="male"
/cell_type="sperm"
/mote=Vector: pELODACL1, Site_1: EcoRI, Site_2: EcoRI

```

Caltech Human BAC Library D"
BASE COUNT 173 a 134 c 162 g 206 t
ORIGIN

Query Match 48.8%; Score 243.8; DB 12; Length 675;
Best Local Similarity 99.2%; Pred. No. 4.5e-65;
Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AACCTGTCGACCTCTATTGACCTATGCTTCGATACGCTTATGAGACATTGTC 60
DB 429 AACCTGTCGACCTCTATTGACCTATGCTTCGATACGCTTATGAGACATTGTC 488
QY 61 AGGCAAAAGATATATATATGCAAACTACGCTTATGATATGATGATGAT 120
DB 489 AGGCAAAAGATATATATATGCAAACTACGCTTATGATATGATGATGAT 548
QY 121 TTGATGTCGACGAGTATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 549 TTGATGTCGACGAGTATGATGATGATGATGATGATGATGATGATGATGAT 608
QY 181 CTTGCTGTCGACCTCTCTCGATGATGATGATGATGATGATGATGATGATGAT 240
DB 609 CTTGCTGTCGACCTCTCTCGATGATGATGATGATGATGATGATGATGATGAT 668
QY 241 GCACAAAC 247
DB 669 GCACAAAC 675

RESULT 2 723 bp DNA linear GSS 29-SEP-2000
AZ344779/c 1M0079E03F Mouse 10kb plasmid UNGCM library Mus musculus genomic
DEFINITION clone UNGCM0079E03 F, DNA sequence.

ACCESSION AZ344779
VERSION AZ344779.1 GI:10424016
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, J.
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL Plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Rm 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0079 row: E column: 03
Seq primer: CGTGTAAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 723.

FEATURES

source
1. 723
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UNGCM0079E03"
/clone_lib="Mouse 10kb plasmid UNGCM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson

BASE COUNT
ORIGIN

Query Match 8.4%; Score 42.2; DB 12; Length 723;
Best Local Similarity 63.1%; Pred. No. 0.052; Mismatches 38; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 2 AACCTGTCGACCTCTATTGACCTATGCTTCGATACGCTTATGAGACATTGTC 61
DB 714 AACCTGTCGACCTCTATTGACCTATGCTTCGATACGCTTATGAGACATTGTC 655
QY 62 GCACAAAGATATATATATGCAAACTACGCTTATGATGATGATGATGAT 104
DB 654 GCACAAAGATATATATATGCAAACTACGCTTATGATGATGATGATGAT 612

RESULT 3 931 bp DNA linear GSS 02-JUN-2000
AZ128097 OSJNB0086H01R CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
LOCUS clone OSJNB0086H01R, DNA sequence.
DEFINITION AZ128097
ACCESSION AZ128097
VERSION AZ128097.1 GI:8204877
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Wing RA
Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Seq primer: GGAACACGATATGACGACG
Class: BAC ends
High quality sequence start: 19
High quality sequence stop: 615.

FEATURES

source
1. 931
Location/Qualifiers
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="OSJNB0086H01R"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/note="Vector: pBACindigo; Site_1: EcoRI; Site_2: EcoRI;
Rice is the most important food crop in the world. Half of

Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1473211419b/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 84.9333 seconds
(without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101a-12_COPY_1_500

Percent score: 500
Sequence: 1 aagcttgatgcacatctattt.....ccctgccgagcgaagaag 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq.032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	8372	22	AD11119
2	33.6	6.7	2811	22	AA20201
3	33.6	6.7	2811	22	AA20201
4	33.4	6.7	2811	22	AA20201
5	33.3	6.6	1214	11	AA004551
6	32.8	6.6	1214	11	AA004551
7	32.6	6.5	1366	21	AA16216
8	32.6	6.5	1366	21	AA16216
9	32.6	6.5	32248	22	AA20412

C	10	32.6	6.5	32248	22	AA137122
C	11	32.6	6.5	32248	22	AA26795
C	12	32.4	6.5	1671	21	AA279977
C	13	32.4	6.5	5003	22	AA339887
C	14	32.4	6.5	5003	22	AA339888
C	15	32.4	6.5	5003	22	AA339888
C	16	32.4	6.5	5003	22	AA339888
C	17	32.4	6.5	5003	22	AA339888
C	18	32.2	6.4	36267	22	AA339888
C	19	32.2	6.4	36267	22	AA339888
C	20	31.8	6.4	480	21	AA339888
C	21	31.8	6.4	480	21	AA339888
C	22	31.8	6.4	480	21	AA339888
C	23	31.8	6.4	480	21	AA339888
C	24	31.8	6.4	480	21	AA339888
C	25	31.8	6.4	480	21	AA339888
C	26	31.8	6.4	480	21	AA339888
C	27	31.8	6.4	480	21	AA339888
C	28	31.6	6.3	6016	23	AA102857
C	29	31.6	6.3	6016	23	AA102857
C	30	31.6	6.3	6016	23	AA102857
C	31	31.4	6.3	17137	24	AA132190
C	32	31.4	6.3	17137	24	AA132190
C	33	31.4	6.3	17137	24	AA132190
C	34	31.2	6.2	893	21	AA60591
C	35	31.2	6.2	893	21	AA60591
C	36	31.2	6.2	1238	22	AAH89875
C	37	31.2	6.2	1238	22	AAH89875
C	38	31.2	6.2	10689	22	AAH89875
C	39	31.2	6.2	10689	22	AAH89875
C	40	31.2	6.2	349980	22	AAH89875
C	41	31.2	6.2	349980	22	AAH89875
C	42	31.2	6.2	251	19	AAH89875
C	43	31.2	6.2	251	19	AAH89875
C	44	31.2	6.2	251	19	AAH89875
C	45	31.2	6.2	251	19	AAH89875

ALIGNMENTS

RESULT 1	AD11119	standard; DNA: 8372 BP.
XX	AD11119	
AC	AD11119	
XX	AD11119	
DT	24-SEP-2001	(first entry)
XX	24-SEP-2001	
DE	Human small cell lung cancer associated gene, SOX21.	
XX	Human small cell lung cancer; therapy: hCAAP; nucleic acid: NA;	
KW	melanoma; cancer: colon; breast; head; neck; transitional cancer;	
KW	leiomyosarcoma; synovial sarcoma; cytostatic; SOX21; ds.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	1167..1997
FT	FT	/*tag- a
FT	FT	/product- "Human SOX21 protein"
PN	W0200153349-A2.	
PD	26-JUL-2001.	
XX	26-JUL-2001.	
XX	19-JAN-2001; 2001MO-USO2015.	
PT	19-JAN-2001; 2001MO-USO2015.	
XX	21-JAN-2000; 200005-0489101.	
XX	21-JAN-2000; 200005-0489101.	
XX	(LUDW-) LUDWIG INST CANCER RES.	
PA	(SLOK) SLOAN KETTERING INST CANCER RES.	
PA	(CORR) CORNELL RES FOUND INC.	

Human musculoskele
Human genomic DNA
Human secreted pro
Genomic sequence #
Genomic sequence #
Human digestive sy
Human digestive sy
Drosophila melanog
Human colon cancer
Human ORF2220
Restriction fragme
5.6 Kb Ell Pseudom
Pseudomonas fluore
E11 genomic DNA fr
5.6 Kb EcoRI-Hind
Human immune/haema
Pseudomonas fluore
P. fluorescens 11k
Drosophila melanog
Human immune syste
Human musculoskele
Human musculoskele
Human secreted pro
PEMY nucleotide se
Rat phosphatidylet
Human bone marrow
C glutamicum codin
Human immune/haema
Human immune/haema
C glutamicum codin
Human biallelic po
Human biallelic po
Human biallelic po
Human biallelic po
Human biallelic po

XX	AAF84712:	
AC		
DF	29-JUN-2001	(first entry)
XX		
DE	Nucleotide sequence of a bacterial glucose-6-phosphate isomerase.	
XX		
KM	Corynebacterium bacteria; glucose-6-phosphate isomerase; gpi; metabolic flux;	
KW	pentose phosphate cycle; L-amino acid; nucleotide production;	
KM	vitamin production; L-amino acid production; ss.	
XX		
OS	Corynebacterium glutamicum.	
XX		
XX	Key	Location/Qualifiers
FT	CDS	373..2025
FT		/tag= a
XX		/product= "glucose-6-phosphate isomerase"
PN	EP1087015-A2.	
XX		
PD	26-MAR-2001.	
XX		
XX	23-AUG-2000: 2000EP-0118052.	
PE		
XX	15-SEP-1999: 99US-0396478.	
XX		
XX	(DEGS) DEGUSSA-HUELS AG.	
PA	(UYNA-) UNIV NAT IRELAND.	
XX		
PI	Dunican LK, McCormack A, Stapleton C, Burke K, O'Donohue M;	
XX	Max A, Moeckel B;	
XX		
DR	WPI: 2001-267593/28.	
DR	P-PSDB: AAB68033.	
XX		
PP	Novel nucleotide sequence from corynebacterium bacteria encoding	
PR	glucose-6-phosphate isomerase, attenuation of which increases metabolic	
PR	flux through pentose phosphate cycle in the bacteria, producing L-amino	
XX	acids	
XX		
PS	Claim 4: Page 13-16; 33pp; English.	
XX		
CC	The present sequence encodes a corynebacterium bacterial glucose-6-phosphate	
CC	isomerase (gpi). After attenuation of the gpi gene, corynebacterium bacteria	
CC	exhibit an improved metabolic flux through the pentose phosphate cycle	
CC	and produce nucleotides, vitamins, and in particular L-amino acids,	
CC	particularly L-lysine and L-tryptophan, in an improved manner.	
CC	Attenuation may be achieved by reducing or switching off either the	
CC	expression of the gpi gene or the catalytic properties of the enzyme	
CC	protein. Microorganisms, in particular Corynebacterium glutamicum in	
CC	which the gpi gene is attenuated or switched off, optionally in	
CC	combination with amplification of further genes, are useful for	
CC	producing nucleotides, vitamins or L-amino acids.	
CC		
XX		
SO	Sequence 2811 BP; 592 A; 744 G; 764 G; 711 T; 0 other;	
XX		
XX	Query Match	6.7%; Score 33.6; DB 22; Length 2811;
XX	Best Local Similarity	50.0%; Pred. No. 3.3;
XX	Matches 84; Conservative	0; Mismatches 84; Indels 0; Gaps 0;
DB	180 GCTGCTCTCCACCCCTCTCGAGTAGCATTTTGGATGTGACAGGAGTCTCCCTCTGG 239	
DB	2804 GGGCGCTCTAGAGTCTCTCGAGAGCTTTTGAATCTTGGATCACCACCATCTGC 2745	
QY	240 GGCACATCAACAAAGAAAGTGTCTAAGGACAAAGAGCGTGGGAAATGACATCTCCA 299	
DB	2744 GGCATTAACCACTATGAGATATACCTCGCAACGAAGTAACCAATCATCAACCCGAGC 2685	
QY	300 TTGGAACAGCCCTGCGCTTACTCATATGCTGAGAGAGCTGTATGGC 347	
DB	2684 CCCACACCAACGAGACTTCCACACAGGAAAGAAACCTTCTGACGGC 2637	

	RESULT	4
ID	AAZ15378	
XX	AAZ15378 standard; cDNA; 774 BP.	
AC	AAZ15378;	
Dn	12-OCT-1999 (first entry)	
XX	Human gene expression product cDNA sequence SEQ ID NO:2847.	
DE		
KW	Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss. 	
OS	Homo sapiens.	
PN	M09338972-AZ.	
PD	05-AUG-1999.	
PF	28-JAN-1999; 99MO-USO1619.	
PR	03-APR-1998; 98US-0080666. 28-JAN-1998; 98US-0072910. 24-FEB-1998; 98US-0075954. 31-MAR-1998; 98US-0080114. 03-APR-1998; 98US-0080515.	
PA	(CHIR-) CHIRON CORP. (HYSE-) HYSEO INC.	
PI	Cikvenjakov R., Dickson M., Drmanac S; Escobedo J., Garcia PD, Garcia V, Giese K., Innis MA, Jones WJ., Kassam A., Kennedy GC., Kita D., Labat I.; Lamsom G., Leshkowitz D., Pot D., Randazzo F., Reinhard C; Stache-Crain B., Suduth-Klinger J., Williams LT; WPt.: 1999-494092/41.	
PT	Novel human genes and their expression products which are differentially expressed in different cell types	
PS	Claim 1; Page 1377; 2479pp; English.	
PP		
XX	The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ1779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5548 polynucleotide sequences given in AAZ12532 to AAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for peptide analogues and antagonists.	
SQ	Sequence 774 BP; 233 A; 119 C; 123 G; 251 T; 48 other:	
	Query Match 6.7%; Score 33.4; DB 20; Length 774;	
	Best Local Similarity 47.8% ; Pred No.1.9; Indels 0;	
	Matches 97; Conservative 0; Mismatches 106; Gaps 0;	
OY	33 TGCAATCAGCGTTTATGGGACAATTGTCTGCAGCAAATAATGAATTCGCAACTCCACG	92

CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAF57303 represent sequences used in the exemplification of the present
CC invention.

CC Sequence 1366 BP; 262 A; 387 C; 327 G; 386 T; 4 other;

Query Match 5.5%; Score 32.6; DB 21; length 1366;
Best Local Similarity 59.6%; Pred. No. 4.8;

Matches 53; Conservative 1; Mismatches 35; Indels 0; Gaps 0;

OY 42 CTTATGGGACATTTGTCAGCAAAAGTATATATGCAACTGACGCTTTATTT 101
DB 275 CTTCTCTTATTTTCTTATGATATGATTTTACATTTTATTCCTTTCTTTT 334
OY 102 TAAATGATTTGGTGTGATTTGATGCTGA 130
DB 335 TGGTTTGATTTGGTTTGGTTGAGGAGA 363

RESULT 9

ABR20412
ID ABR20412 standard; DNA; 32248 BP.

AC ABR20412;

XX 23-JAN-2002 (first entry)

DE Human nervous system related polynucleotide SEQ ID NO 12743.

XX Human; neoplastic; neuroprotective; cytostatic; dermatological; vitruclide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; antiviral;
KW antiparasitic; antitickling; antianemic; antitubercular; cancer;
KW antihemematic; hepatotropic; cerebroprotective; antihemematic;
KW antileptogenic; antidiabetic; antileptogenic; antileptogenic;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US01334.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217487.
PR 26-JUL-2000; 2000US-0218299.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225457.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232968.
PR 14-SEP-2000; 2000US-0232997.
PR 14-SEP-2000; 2000US-0232998.
PR 14-SEP-2000; 2000US-0232999.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241807.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0242417.
PR 08-NOV-2000; 2000US-0242474.
PR 08-NOV-2000; 2000US-0242475.
PR 08-NOV-2000; 2000US-0242476.
PR 08-NOV-2000; 2000US-0242477.
PR 08-NOV-2000; 2000US-0242478.
PR 08-NOV-2000; 2000US-0245233.
PR 08-NOV-2000; 2000US-0245234.
PR 08-NOV-2000; 2000US-0245235.
PR 08-NOV-2000; 2000US-0245236.
PR 08-NOV-2000; 2000US-0245237.

17-NOV-2000; 2000US-0249218-
PR 17-NOV-2000; 2000US-0249244-
PR 17-NOV-2000; 2000US-0249245-
PR 17-NOV-2000; 2000US-0249264-
PR 17-NOV-2000; 2000US-0249265-
PR 17-NOV-2000; 2000US-0249297-
PR 17-NOV-2000; 2000US-0249299-
PR 17-NOV-2000; 2000US-0249300-
PR 01-DEC-2000; 2000US-0250160-
PR 01-DEC-2000; 2000US-0250391-
PR 05-DEC-2000; 2000US-0251030-
PR 05-DEC-2000; 2000US-0251988-
PR 06-DEC-2000; 2000US-0256119-
PR 06-DEC-2000; 2000US-0251479-
PR 08-DEC-2000; 2000US-0251856-
PR 08-DEC-2000; 2000US-0251869-
PR 08-DEC-2000; 2000US-0251899-
PR 08-DEC-2000; 2000US-0251990-
PR 11-DEC-2000; 2000US-0254097-
PR 05-JAN-2001; 2001US-0259678-
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX Disclosure; SEQ ID NO 1769; 980bp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are
CC also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence is a genomic DNA encoding a partial novel secreted protein of
CC the invention.

Query Match 6.5%; Score 32.6; DB 22; Length 32248;
Best Local Similarity 63.3%; Pred. No. 27;
Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 43 TTTATGGACATTTGTCAAGCAAGATATAATTAATGCAAACTCAGCCCTTTATTTT 102
DB 25792 TCTAAGGCTCTGTGCTGAGAAAGATATGATGATCAATTTGTGCTTTCTCT 25851
OY 103 AATTAAGTTGGTGTGATT 121
DB 25852 CAATCTGCTTTGTGCACT 25870

RESULT 12
AAC79977/c
ID AAC79977 standard; cDNA; 1671 BP.
XX
XX AAC79977;
AC
XX 12-FEB-2001 (first entry)
DT
XX
XX Human secreted protein encoding cDNA for gene 30.
DE
XX
XX Secreted protein; human; immunosuppressive; antiarthritic; antihemetic;
KW antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;
KW neurologic; neuroprotective; antibacterial; vitruide; fungicide;
KW ophthalmological; vulnery; gene therapy; treatment; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; ocular disorder;
KW cerebrovascular disorder; nervous system disorder; infection; skin aging;
KW wound healing; epithelial cell proliferation; transplantation; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200058357-A1.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 23-MAR-2000; 2000MO-US07723.
PF
XX
XX 26-MAR-1999; 99US-0126506.
PR
XX
XX 07-JAN-2000; 2000US-0174852.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-611704/58.
XX
XX P-PSDB; AAB45054.
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
XX Claim 1a; Page 357; 418bp; English.
XX
XX This invention describes novel isolated nucleic acid molecules (I)
CC encoding a human secreted protein (II) which have immunosuppressive,
CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiac,
CC vasotropic, cerebroprotective, neurologic, neuroprotective, antibacterial,
CC vitruide, fungicide, ophthalmological and vulnery activity and can be
CC used for gene therapy. (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to (II) can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC disorders which are diagnosed or treated include autoimmune diseases
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi and ocular disorders e.g. corneal infection. The
CC polypeptides can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues and in chemotaxis. The polypeptides can also be
CC used as a food additive or preservative to increase or decrease storage
CC capabilities.
CC
XX
XX Sequence 1671 BP; 481 A; 294 C; 325 G; 571 T; 0 other;

Query Match 6.5%; Score 32.4; DB 21; Length 1671;
Best Local Similarity 56.6%; Pred. No. 6.2;
Matches 60; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
OY 17 ATTTGGACTATGCTTCGACGACTTATGAGAACATTTGTCAGCAAGATATATA 76
    |||||
DB 1198 ATTTGAGCTATGCTTCGACGACTTATGAGAACATTTGTCAGCAAGATATATA 1139
    |||||
OY 77 ATGCGCAACTGACGCTTTTATTTAATTTGATTTGCTGATTT 122
    |||||
DB 1138 ACAGGACAGCTACTCTTAACTGCTTTGCTTTGAGTACT 1093
    |||||

RESULT 13
AAS39887/c
ID AAS39887 standard; DNA, 5003 BP.
AC AAS39887;
XX
XX 17-DEC-2001 (first entry)
DE Genomic sequence #306 encoding human colon associated polypeptide.
XX
XX Human; colon cancer; congenital abnormality; infection; colitis;
KM Inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KM Intestinal inflammatory disorder; malabsorption syndrome; gastritis;
KM sigmoid disease; antibacterial; antiviral; antiinflammatory;
KM cystostatic; ds.
XX
XX Homo sapiens.
XX
XX NC020015302-A2.
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01240.
FE
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198113.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0230963.
PR 14-AUG-2000; 2000US-0234518.
PR 14-AUG-2000; 2000US-0234519.
PR 14-AUG-2000; 2000US-0235214.
PR 14-AUG-2000; 2000US-0235256.
PR 14-AUG-2000; 2000US-0235257.
PR 14-AUG-2000; 2000US-0235270.
PR 14-AUG-2000; 2000US-0235276.
PR 14-AUG-2000; 2000US-0235277.
PR 14-AUG-2000; 2000US-0235279.
PR 18-AUG-2000; 2000US-0235758.
PR 22-AUG-2000; 2000US-0236681.
PR 22-AUG-2000; 2000US-0236681.
PR 22-AUG-2000; 2000US-0237182.
PR 23-AUG-2000; 2000US-0237009.
PR 30-AUG-2000; 2000US-0238924.
PR 01-SEP-2000; 2000US-0239287.
PR 01-SEP-2000; 2000US-0239343.
PR 01-SEP-2000; 2000US-0239344.
PR 01-SEP-2000; 2000US-0239345.

PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231245.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 20-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240560.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241185.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0246177.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
```


Fri Oct 11 09:30:35 2002

us-09-489-101a-11_copy_1_500.rni

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 10, 2002, 12:39:37 ; Search time 18.733 Seconds
(without alignments)
6536.054 Million cell updates/sec

Title: US-09-489-101a-11_COPY_1_500
Perfect score: 1 tgcgcgagggatcgagcgc.....ccgagggctcctcgtcattt 500
Sequence: 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2_6/prodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCFOS.COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	32.2	6.4	1556	4 US-09-043-937A-3	Sequence 3, Appl1
2	31.6	6.3	311	3 US-08-937-580-20	Sequence 20, Appl1
3	31.6	6.3	311	4 US-09-336-039-20	Sequence 20, Appl1
4	30	6.0	311	3 US-08-937-580-19	Sequence 19, Appl1
5	30	6.0	311	4 US-09-336-039-19	Sequence 19, Appl1
6	30	6.0	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
7	29.4	5.9	1548	4 US-08-379-580-3	Sequence 3, Appl1
8	29.4	5.9	4307	1 US-08-190-687B-7	Sequence 156, App
9	29.4	5.9	13254	1 US-08-276-852-156	Sequence 170, App
10	29.4	5.9	13254	1 US-08-899-575-156	Sequence 156, App
11	29.4	5.9	13254	1 US-08-899-575-170	Sequence 170, App
12	29.4	5.9	13254	1 US-08-899-575-156	Sequence 156, App
13	29.4	5.9	13254	1 US-08-899-575-170	Sequence 170, App
14	29.4	5.9	13254	1 US-08-899-575-156	Sequence 156, App
15	29.4	5.9	13254	1 US-08-899-575-170	Sequence 170, App
16	29.4	5.9	13254	1 US-08-899-575-156	Sequence 156, App
17	29.4	5.9	13254	1 US-08-899-575-170	Sequence 170, App
18	29.4	5.9	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
19	29.4	5.9	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
20	28.8	5.8	1784	2 US-08-998-416-1131	Sequence 1131, Ap
21	28.8	5.8	1784	3 US-08-808-931-19	Sequence 19, Appl1
22	28.8	5.8	1784	3 US-08-808-323-19	Sequence 19, Appl1
23	28.8	5.8	1784	3 US-09-050-603A-19	Sequence 19, Appl1
24	28.8	5.8	1784	3 US-09-102-420B-19	Sequence 19, Appl1
25	28.6	5.7	380	4 US-09-497-698-19	Sequence 19, Appl1
26	28.6	5.7	380	4 US-08-776-971-114	Sequence 114, App
27	28.4	5.7	2943	1 US-08-042-747A-7	Sequence 7, Appl1
28	28.4	5.7	2826	3 US-09-010-998-1	Sequence 1, Appl1

28	28.4	5.7	7218	1 US-08-232-463-14	Sequence 14, Appl1
29	28.2	5.6	1379	2 US-08-437-607A-1	Sequence 1, Appl1
30	28.2	5.6	1380	2 US-08-437-607A-4	Sequence 4, Appl1
31	28.2	5.6	2344	3 US-08-893-852A-2	Sequence 2, Appl1
32	28.2	5.6	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
33	28	5.6	423	1 US-08-470-179-109	Sequence 109, Appl1
34	28	5.6	1537	6 5504194-1	Patent No. 5504194
35	28	5.6	7201	4 US-08-815-088-3	Sequence 3, Appl1
36	27.8	5.6	12001	4 US-08-458-568A-11	Sequence 11, Appl1
37	27.6	5.5	1060	2 US-08-385-132A-14	Sequence 14, Appl1
38	27.6	5.5	1056	3 US-08-825-168-17	Sequence 17, Appl1
39	27.6	5.5	1636	3 US-08-825-227-17	Sequence 17, Appl1
40	27.6	5.5	1636	3 US-08-653-241-17	Sequence 17, Appl1
41	27.6	5.5	1636	4 US-08-398-526-17	Sequence 17, Appl1
42	27.6	5.5	2241	3 US-09-144-759-17	Sequence 17, Appl1
43	27.6	5.5	2295	3 US-09-144-759-19	Sequence 19, Appl1
44	27.6	5.5	2301	3 US-09-144-759-21	Sequence 21, Appl1
45	27.6	5.5	2676	3 US-09-212-971-11	Sequence 11, Appl1

ALIGNMENTS

RESULT 1
US-09-043-937A-3
Sequence 3, Application US/09043937A
Patent No. 6211432

GENERAL INFORMATION:
APPLICANT: BOUDIER, ALAIN-MICHEL
PICHON, MAGALIE
GRIMA-DENNETT, JACQUELINE
BRICKERT, MICHEL
GAMAS, PASCAL
BRIAT, JEAN-FRANCOIS

TITLE OF INVENTION: DNA SEQUENCES CODING FOR CINNAMYL-CAA
REDUCTASE, AND APPLICATIONS THEREOF IN THE CONTROL OF
LIGNIN CONTENTS IN PLANTS

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE, P.C.
STREET: 110 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/043.937A
FILING DATE: 24-Jul-1998
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/01544
FILING DATE: 03-OCT-1996
APPLICATION NUMBER: FR 95.11623
FILING DATE: 03-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B J
REGISTRATION NUMBER: 36.663
REFERENCE/DOCKET NUMBER: 1487-20

TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4100
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1556 base pairs
TYPE: nucleic acid
TOPOLOGY: double
MOLECULE TYPE: linear

MOLECULE TYPE: cDNA to mRNA

Fri Oct 11 09:30:35 2002

us-09-489-101a-11_copy_1_500.rni

Page 2

FEATURE:
NAME/KEY: CDS
LOCATION: 195, 1310
SEQUENCE DESCRIPTION: SEQ ID NO: 3
US-09-043-937A-3

Query Match
Best Local Similarity 45.2%; Score 32.2; DB 4; Length 1556;
Matches 118; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 238 GGGGCGGGGGGACGAAAGGCTCCCGAATTCTTTTTCACGACGAGCCGAGCG 297
DB 478 GCGCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537
QY 298 GCGCGGTAATGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 357
DB 538 AGCAAAATGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597
QY 358 GTTCCGAGGGGGGACGAAAGGCTCCCGAATTCTTTTTCACGAGGATCCGAGGAT 417
DB 598 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 657
QY 418 AGCTAGAGCCGACGAGAGCTGTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 477
DB 658 AGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 717
QY 478 AGCCCGGCGGCTTCCTGCTGAT 498
DB 718 AAACGAGAACTGTGCTGCT 738

RESULT 2

US-08-937-580-20/c
Sequence 20, Application US/08937580
Patent No. 6013510
GENERAL INFORMATION:
APPLICANT: Harris, James M.
TITLE OF INVENTION: Identification of a DNA Region
TITLE OF INVENTION: Potentially Useful for the Detection of Mycobacterium
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Becton Dickinson and Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: New Jersey
COUNTRY: USA
ZIP: 07417-6800
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,580
FILING DATE: 25-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3690/5510-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-847-7166
TELEFAX: 201-848-9228
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORIGINAL SOURCE:

ORGANISM: Mycobacterium gastr

US-08-937-580-20
Query Match
Best Local Similarity 53.2%; Score 31.6; DB 3; Length 311;
Matches 67; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 203 CCTAAGTACTTACACGCTCAGTACGAGGCGGCGGCGGCGGCGGCGGCGGCGG 262
DB 148 CCGTAGCATTCGACCGCGGATGCGATCAGAGCGGCGGCGGCGGCGGCGGCGGCGG 89
QY 263 CCGAATTTTCTTTCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 322
DB 88 ACCGAGTTGTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 29
QY 323 CATGAC 328
DB 28 CATGCG 23

RESULT 3

US-09-336-039-20/c
Sequence 20, Application US/09336039
Patent No. 6291176
GENERAL INFORMATION:
APPLICANT: Harris, James M.
TITLE OF INVENTION: Identification of a DNA Region
TITLE OF INVENTION: Potentially Useful for the Detection of Mycobacterium
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Becton Dickinson and Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: New Jersey
COUNTRY: USA
ZIP: 07417-6800
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/336,039
FILING DATE: 18-Jun-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,580
FILING DATE: 25-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3690/5510-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-847-7166
TELEFAX: 201-848-9228
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Mycobacterium gastr
US-09-336-039-20
Query Match
Best Local Similarity 53.2%; Score 31.6; DB 4; Length 311;
Matches 67; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Db 148 CCGTAGGCAATTCGACCGGATGCGATCAGAGCAGCTGGACGACGACGCGTT 89
OY 263 CCCGAACCTTTTTCAGCCAGCCGACGAGGCGCTCGTATGATGGCCAGGCG 322
Db 88 ACCCGAGTTCGTCGCGCCACGACGCGACCCGTCGATGCTGACGATGCTGGCGG 29
OY 323 CATCAC 328
Db 28 CATCGC 23

RESULT 4

US-08-937-580-19/c
Sequence 19, Application US/08937580
Patent No. 6013510
GENERAL INFORMATION:
APPLICANT: Harris, James M.
TITLE OF INVENTION: Identification of a DNA Region
TITLE OF INVENTION: Potentially Useful for the Detection of Mycobacterium
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Becton Dickinson and Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: New Jersey
COUNTRY: USA
ZIP: 07417-6800
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,580
FILING DATE: 25-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3690/5510-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-847-7166
TELEFAX: 201-848-9228
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Mycobacterium gastr1 and Mycobacterium
US-08-937-580-19

Query Match 6.0%; Score 30; DB 3; Length 311;
Best Local Similarity 47.6%; Pred. No. 1.2;
Matches 60; Conservative 8; Mismatches 58; Indels 0; Gaps 0;

OY 203 CCTAAGTAACCTTACCACTGCTAGGCAAGAGGAGGCGTGGGTGAACGAAGGCT 262
Db 148 CCGTAGGCAATTCGACCGGATGCGATCAGAGCAGCTGGACGACGACGCGTT 89
OY 263 CCCGAACCTTTTTCAGCCAGCCGACGAGGCGCTCGTATGATGGCCAGGCG 322
Db 88 ACCCGAGTTCGTCGCGCCACGACGCGACCCGTCGATGCTGACGATGCTGGCGG 29
OY 323 CATCAC 328
Db 28 CATCGC 23

RESULT 5
US-09-336-039-19/c
Sequence 19, Application US/09336039
Patent No. 6291176
GENERAL INFORMATION:
APPLICANT: Harris, James M.
TITLE OF INVENTION: Identification of a DNA Region
TITLE OF INVENTION: Potentially Useful for the Detection of Mycobacterium
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Becton Dickinson and Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: New Jersey
COUNTRY: USA
ZIP: 07417-6800
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/336,039
FILING DATE: 18-Jun-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,580
FILING DATE: 25-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3690/5510-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-847-7166
TELEFAX: 201-848-9228
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Mycobacterium gastr1 and Mycobacterium
US-09-336-039-19

Query Match 6.0%; Score 30; DB 4; Length 311;
Best Local Similarity 47.8%; Pred. No. 1.2;
Matches 60; Conservative 8; Mismatches 58; Indels 0; Gaps 0;

OY 203 CCTAAGTAACCTTACCACTGCTAGGCAAGAGGAGGCGTGGGTGAACGAAGGCT 262
Db 148 CCGTAGGCAATTCGACCGGATGCGATCAGAGCAGCTGGACGACGACGCGTT 89
OY 263 CCCGAACCTTTTTCAGCCAGCCGACGAGGCGCTCGTATGATGGCCAGGCG 322
Db 88 ACCCGAGTTCGTCGCGCCACGACGCGACCCGTCGATGCTGACGATGCTGGCGG 29
OY 323 CATCAC 328
Db 28 CATCGC 23
RESULT 6
US-09-103-840A-1
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:

```

1  APPLICANT: FLEISCHMAN, Robert D.
2  APPLICANT: WHITE, Owen R.
3  APPLICANT: FRASER, Claire M.
4  APPLICANT: VENTER, John C.
5  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
6  TITLE OF INVENTION: TUBERCULOSIS
7  FILE REFERENCE: 24366-20007.00
8  CURRENT APPLICATION NUMBER: US/09/103,840A
9  CURRENT FILING DATE: 1998-06-24
10 NUMBER OF SEQ ID NOS: 2
11 SOFTWARE: PatentIn Ver. 2.1
12 SEQ ID NO 1
13 LENGTH: 4411529
14 TYPE: DNA
15 ORGANISM: Mycobacterium tuberculosis
16 OTHER INFORMATION: H37Rv
17 US-09-103-840A-1

```

Query Match	6.0%	Score 30;	DB 4;	Length 4411529;
Best Local Similarity	52.4%	Pred. No. 47;		
Matches	66;	Conservative	0;	Mismatches 60;
			Indels	0;
			Gaps	0;
Oy	316	CAGGGGCGCATCTACATGCGAACCTGTCAATCACGGGCTCCGGGTTCGAGGGGCGGACCA	375	
Db	2151568	CAGGATCCGACCCCGGTAATCTGTCTAATACATCCGCCCTCCGCAAGCGGTGTGTTATGCG	2151627	
Oy	376	AGCCGCAACCCCGGGGAATCTCCAGGACAGGTATTTAAGGGGCCCAAGCTAAGACCCCAAGGACGA	435	
Db	2151628	GGCGAGCAACCTTTCAGAGATCCCTCGGTTATCTCCGGCTCTGTGGGTGGGGTCCACCTGCG	2151687	
Oy	436	CTGTGCA	441	
Db	2151688	CAGGGA	2151693	

RESULT 7
US-08-379-580-3
Sequence 3, Application US/08379580
Patent No. 6180362
GENERAL INFORMATION:
APPLICANT: Duchesne, Marc
APPLICANT: Schweighofer, Fabien
TITLE OF INVENTION: Peptides Inhibiting Ras Protein
TITLE OF INVENTION: Activity, Preparation and Use Thereof
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/379,580
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00772
FILING DATE: 28-Jul-1993
PRIOR APPLICATION DATA: FR 92/09433
FILING DATE: 03-Jul-1992
ATTORNEY/AGENT INFORMATION:
NAME: SMITH, JULIE K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: S792049-US
TELECOMMUNICATION INFORMATION:

? TELEPHONE: (610)454-3839
 ? TELEFAX: (610)454-3808
 ? INFORMATION FOR SEQ ID NO: 3:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 1548 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: cDNA
 ? ORIGINAL SOURCE:
 ? ORGANISM: Homo sapiens
 ? FEATURE:
 ? NAME/KEY: CDS
 ? LOCATION: 1..1548
 ? US-08-379-580-3

Query Match %: 5.9%; Score 29.4; DB 4; Length 1548;
Best Local Similarity 70.9%; Pled. No. 4.1;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0.

```
OY      68 cccctcctcccccccacatcgcttggcgtgagcgagcgsgtcggsgggg 122  
        |||||  
Db       14 cccctctgcgcccttcgcccttactctcccccctttggggggggccatcgaggacagt 68
```

US-08-190-687B-7
Sequence 7, Application US/08190687B
Patent No. 5760203
GENERAL INFORMATION:
APPLICANT: Wong, Gail L.
APPLICANT: Martin, George
APPLICANT: McCormick, Francis P.
APPLICANT: Rubinfield, Bonnie
APPLICANT: O'Rourke, Edward C.
APPLICANT: Clark, Robin
TITLE OF INVENTION: GAP Gene Sequences
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,687B
FILING DATE: 02-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/774,644
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/260,807
FILING DATE: 21-OCT-1988
APPLICATION NUMBER: 07/230,761
FILING DATE: 10-AUG-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 27527/31898
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 4307 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 119..3259
US-08-190-687B-7

Query Match 5.9%; Score 29.4; DB 1; Length 4307;
Best Local Similarity 70.9%; Pred. No. 6.6;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 68 CCCCTGCGCCCGCCCAATCTGCTGCGGGGGTGGGGGGGT 122
Db 522 CCCCTGCGCCCGCCCAATCTGCTGCGGGGGTGGGGGGGT 576

RESULT 9

US-08-276-852-156/c
Sequence 156, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop 1PC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ. ID NO.: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-276-852-156

Query Match 5.9%; Score 29.4; DB 1; Length 13254;
Best Local Similarity 50.3%; Pred. No. 11;
Matches 72; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 116 GGGGGTACCTCTCAGGTTTCCTTGAACCTTTGAAACCTAATGCGCT 175
Db 8968 GGGGAGCATGACCTCAGCATTTGCTGTAATCTTGACCCAGATGACAGCGCC 8909
Qy 176 CTGAGTGGCCCTGCTGAGCTCCCGCTCTTAAGTACTTTACACAGTACAGGCGAA 235
Db 8908 GGTAGTACCTCAGCATATCTCTGCAATGCTTGTGCGCCACACACAGTAAAT 8849
Qy 236 GAGGGCTGGGGTGAACGAAG 258
Db 8848 ACGGACCTGGGGCCAGGAAG 8826

RESULT 10

US-08-276-852-170
Sequence 170, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop 1PC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ. ID NO.: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-276-852-170

Query Match 5.9%; Score 29.4; DB 1; Length 13254;
Best Local Similarity 50.3%; Pred. No. 11;
Matches 72; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 116 GGGGGTACCTCTCAGGTTTCCTTGAACCTTTGAAACCTAATGCGCT 175
Db 4287 GGGGAGCATGACCTCAGCATTTGCTGTAATCTTGACCCAGATGACAGCGCC 4346

GENERAL INFORMATION:

APPLICANT: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08743
FILING DATE: 11-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
PCT-US95-08743-156

Query Match

5.9%; Score 29.4; DB 5; Length 13254;
Best Local Similarity 50.3%; Pred. No. 11;

Matches 72; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 116 GGGGGGTACCTCTCAGGCTTCTTCAAACTTTTGAACCCCAATTGGTGCGCT 175
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8968 GGGCAGCAGCATGACCTCAGCATTTGTCCTGTAATCTTGACCCCGACATACAGCAGCGC 8909
QY 176 CTGAGTGGGCGCTCGGAGCTCCCGCTCCTAAGTAACTTACCAAGTACACTAGGCCAA 235
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8908 GGTAGTGAACCTTCACAGATATCCCTGCCATAGGCTTTGTCTGCGCCACACACAGTAAT 8849
QY 236 GAGGGCGTGGGGGTGACGAAG 258
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8848 ACGGACCTTGGGGGCCCGAGGAAG 8826

Search completed: October 10, 2002, 18:44:40
Job time : 1790.73 secs

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: <http://genome-gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagai, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matahiki, M., Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T., Matsura
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.

Downloaded from ascelibrary.org by New York University on 06/01/15. Copyright ASCE, For All Rights Reserved, No part of this document may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage or retrieval system, without permission in writing from ASCE.

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:455568"
/clone_id="Scars mouse placenta 4NbMp3.5 14.5"
/sex="unkn"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="pM10B"
/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
       a modified polylinker; Site.1: Not I; Site.2: Eco RI"

```

	BASE COUNT	116 a	128 c	154 g	237 t	142 others
ORIGIN						
Query Match			7.5%;	Score 37.4;	DB 12;	Length 777;
Best Local Similarity			46.7%;	Pred. No. 12;		
Matches	49;	Conservative	19;	Mismatches	37;	Indels 0; Gaps 0;
OY	42	AAGTGTGTTTCTAAATTCGGAAGGCCCTCTGCCTCCCGCCCAATCTTGCGC	101			
I::I::	I::I::	I::I::	I::I::	I::I::	I::I::	I::I::
Dd	746	ARAKTTTGGCCAMAGCTGDSGGACCCCCCCCCCCCCCCCCCCCCCGGGGG	687			

[illegible][illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 84.9333 Seconds
(without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101A-11_COPY_1_500
Perfect score: 500
Sequence: 1 tggccgggggagatgggagcc.....ccggcggtctctcgtgattt 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_032802.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	500	100.0	2509	22	AAD11118	Human small cell 1
2	121.8	24.4	2378	23	AAS87738	DNA encoding novel
3	41.8	8.4	506	22	AA181822	Human polynucleoti
4	38.2	7.6	1468	22	ABA05613	Human bone marrow
5	36.2	7.2	379	22	AA192147	Human polyomucleoti
6	35.4	7.1	114955	20	AA53491	Human adenocarcinoma A1
7	34.4	6.9	686	22	AA552973	Human immune/haema
8	34.2	6.8	637	22	AA564084	Human prostate CDN
9	32.6	6.5	1714	24	AB199331	Mouse ischaemic co

10	32.4	6.5	554	22	AA116574	Human breast cancer
11	32.4	6.5	16106	22	AAK83468	Human immune/haema
12	32.4	6.5	16161	22	AAK83469	Human immune/haema
13	32.2	6.4	1556	18	AA194535	Maize cinnamoyl Co
14	32.2	6.4	1177	22	AAH94477	Human foetal CDNA,
15	31.8	6.4	1623	22	AAAF30438	Rice apoptosis ind
16	31.6	6.3	994	22	AA115399	Human breast cancer
17	31.2	6.2	646	22	AA564368	DNA encoding novel
18	31.2	6.2	12710	22	ABA14807	Human nervous syst
19	31.2	6.2	985	22	AA115317	Human breast cancer
20	31.1	6.2	21888	22	AAK73863	Human immune/haema
21	30.8	6.2	187	21	AAK11389	Human secreted pro
22	30.8	6.2	387	22	AA559169	Human cancer relat
23	30.8	6.2	1950	20	AAH84531	Human secreted pro
24	30.8	6.2	1950	22	AAH83114	Human secreted pro
25	30.8	6.2	2033	21	AAA08502	Human ISRE-binding
26	30.8	6.2	2476	22	AA531013	Human diagnostic a
27	30.8	6.2	4500	16	AA086237	Megakaryocyte stem
28	30.8	6.2	4548	16	AA086237	Human immune/haema
29	30.8	6.2	4758	22	AAK83830	Human nervous syst
30	30.8	6.2	16329	22	AA114411	Human immune/haema
31	30.6	6.1	1559	20	AAK25211	Maize cinnamoyl-Co
32	30.6	6.1	2202	22	AAK08650	Human N-copine hom
33	30.6	6.1	3481	23	AB121297	Drosophila melanog
34	30.6	6.1	11470	23	AB121296	Drosophila melanog
35	30.6	6.1	14175	22	AA527814	DNA encoding novel
36	30.6	6.1	14175	22	AAK78859	Human immune/haema
37	30.4	6.1	1278	16	AA080198	Klebsiella sp. nit
38	30.4	6.1	6779	22	AA559570	Protonibacterium
39	30.2	6.0	2568	6	AA50493	Sequence of skc st
40	30.2	6.0	20445	24	AA519905	Human cDNA clone (
41	30.2	6.0	811	22	AAH06402	Human protein HPO3
42	30	6.0	1520	22	AAH68564	Drosophila melanog
43	30	6.0	2078	23	AB111613	Drosophila melanog
44	30	6.0	2798	22	AAH16206	Human cDNA sequenc
45	30	6.0	4216	23	AB111612	Drosophila melanog

ALIGNMENTS

RESULT 1	
ID	AAD11118
AD11118	standard; DNA: 2509 BP.
XX	
AC	AAD11118:
XX	
DT	24-SEP-2001 (first entry)
XX	
DE	Human small cell lung cancer associated gene, SOX3.
XX	
KW	Human; small cell lung cancer; therapy: hCAP; nucleic acid; NA;
KW	melanoma; cancer; colon; breast; head; neck; transitional cancer;
KW	leiomyosarcoma; synovial sarcoma; cytostatic; SOX3; ds.
XX	
OS	Homo sapiens.
XX	
EH	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "Human SOX3 protein"
PN	WO20015349-A2.
XX	
PD	26-JUL-2001.
XX	
PF	19-JAN-2001; 2001WO-US02015.
XX	
PR	21-JAN-2000; 2000US-0489101.
XX	
PA	(LUDWIG) LUDWIG INST CANCER RES.
PA	(SLOK) SLOAN KETTERING INST CANCER RES.
PA	(CORR) CORNELL RES FOUND INC.

XX Stoeckert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
XX
XX WPI: 2001-457597/49.
DR P-PSDB; AAE05813.

XX Isolated polypeptide, used to treat or prognose a disorder
PT characterized by expression of a hCAP e.g. cancer, is encoded by an
PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -

PS Claim 57; Page 98-99; 152pp; English.

XX The invention relates to nucleic acids and encoded polypeptides which
CC are cancer associated antigens expressed in patients afflicted with
CC small cell lung cancer. The molecules provided by the invention can be
CC used in the diagnosis, monitoring, research or treatment of conditions
CC characterised by the expression of one or more cancer associated
CC antigens. The polypeptide is used to treat a disorder characterised by
CC expression of a hCAP, and determine regression, progression or onset
CC of a condition characterised by expression of an abnormal amount of a
CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
CC are small and non-small cell lung cancer, melanoma, colon, breast, head
CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
CC The present sequence is a small cell lung cancer associated gene
CC designated as NY-SCLC-9, encoding human SOX3 protein.

SQ Sequence 2509 BP; 456 A; 818 C; 757 G; 478 T; 0 other;

Query Match 100.0%; Score 500; DB 22; Length 2509;

Best Local Similarity 100.0%; Pred. No. 2.9e-143;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGCGGGGATGGGGGCGGCTGCTGACAGGCTGCAAAATGTTCTTAATC 60
DB 1 TGGCGGGGATGGGGGCGGCTGCTGACAGGCTGCAAAATGTTCTTAATC 60
QY 61 CGAAGCGCCCTCGCCCTCCGCCCAATCTGCTGCGGGGCTGGGGGGG 120
DB 61 CGAAGCGCCCTCGCCCTCCGCCCAATCTGCTGCGGGGCTGGGGGGG 120
QY 121 GTCACCTCTCAGAGTTCCTTCTTCAAACTTTTGAACCTTAATGGTGGCTCTGAG 180
DB 121 GTCACCTCTCAGAGTTCCTTCTTCAAACTTTTGAACCTTAATGGTGGCTCTGAG 180
QY 181 TGGGCTCGGAGCTCCGCTCTTAAGTACCTTACAGCTCCTAGGCCAAGAGG 240
DB 181 TGGGCTCGGAGCTCCGCTCTTAAGTACCTTACAGCTCCTAGGCCAAGAGG 240
QY 241 GCGTGGGGTGAAGAGGCTCCCAACTTTTTTTCAGCCAGGCCGAGGGGG 300
DB 241 GCGTGGGGTGAAGAGGCTCCCAACTTTTTTTCAGCCAGGCCGAGGGGG 300
QY 301 TCGGTAAATGATTTGGCCAGGCGCATCATCGCAACTGTCAATCAAGGGTCTCCGGGTT 360
DB 301 TCGGTAAATGATTTGGCCAGGCGCATCATCGCAACTGTCAATCAAGGGTCTCCGGGTT 360
QY 361 GCGAGGGGCGGACCAAGCCCAAGCCCGGGGAATCCGAGGATTAATAAGGGGCCAGC 420
DB 361 GCGAGGGGCGGACCAAGCCCAAGCCCGGGGAATCCGAGGATTAATAAGGGGCCAGC 420
QY 421 TAGAGCCAGGCAAGCTGATGAGACTGTTCGAGAGACTCATCAGTCCGAGAGC 480
DB 421 TAGAGCCAGGCAAGCTGATGAGACTGTTCGAGAGACTCATCAGTCCGAGAGC 480
QY 481 CCGGCGGTTCTCTGATTT 500
DB 481 CCGGCGGTTCTCTGATTT 500

RESULT 2

AA887738
ID AA887738 standard; cDNA: 2378 BP.
XX

AC AA887738;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #23542.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR P-PSDB; ABG23551.

PT New isolated polynucleotide and encoded polypeptides, useful in

PS Claim 1; SEQ ID No 23542; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 2378 BP; 590 A; 708 C; 664 G; 416 T; 0 other;

Query Match 24.4%; Score 121.8; DB 23; Length 2378;

Best Local Similarity 98.4%; Pred. No. 3.2e-27;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 279 TCCAGCAGGCGGAAGGCGGCTCGGTAAATGATGGCCAGGCGCATCTGCAACTTG 338
DB 444 TGTGTGCCAGGCCGAAGCGGGGCTCGGTAAATGATGGCCAGGCGCATCTGCAACTTG 503
QY 339 TCATACAGGCTCTCGGGTTCGAGGGGAGCAAGCCCAAGCCCGGGGAATCCGA 398
DB 504 TCATACAGGCTCTCGGGTTCGAGGGGAGCAAGCCCAAGCCCGGGGAATCCGA 563
QY 399 GCAAG 403
DB 564 GCAAG 568

AAI92147
ID AAI92147 standard; cDNA; 379 BP.
XX
AC AAI92147;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 12207.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PE 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RJ;
XX
DR WPI; 2001-514838/56.
XX
DR P-PSDB; AAO12216.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 12207; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI92147-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/publicated_pat_sequences](http://wipo.int/pub/publicated_pat_sequences).
XX
SQ Sequence 379 BP; 99 A; 97 C; 81 G; 102 T; 0 other;
Query Match 7.2%; Score 36.2; DB 22; Length 379;
Best Local Similarity 56.2%; Pred. No. 0.32;
Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
Y 55 AATTCGAGGCGCCCTGCGCCCTCCCAATCTCTCGCTCGCGGGGTGGGGGT 114
D 245 AATTCGAGGCGCGGACACCCCGCCCGCCCGCAAAACCTCTCTCGGGGTGCTCTGG 304
Y 115 GGGGGGTCACCTCTCAGCTTCGTTCTTTGAATTTGAACCCATATTGGTGGCC 174
D 305 GGGGCGCACCGCTTGGTGTGTTTTTTCGCAAAATTTTAAAGTCCGGGTTTTCGCC 364
Y 175 T 175
D 365 T 365
RESULT 6

AA53491/C
ID AA53491 standard; DNA; 114955 BP.
XX
AC AA53491;
XX
DT 05-JUL-1999 (first entry)
XX
DE Human adenosine A1 receptor antisense oligonucleotide fragment.
XX
KW Antisense oligonucleotide; multiple target; antisense treatment;
KW impaired respiration; inflammation; lung disease;
KW pulmonary vasoconstriction; inflammation; allergic rhinitis;
KW acute asthma; allergy; asthma; impeded respiration;
KW respiratory distress syndrome; pain; cystic fibrosis;
KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;
KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
KW colon cancer; breast cancer; lung cancer; pancreatic cancer;
KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
KW prostate cancer; ss.
XX
OS Synthetic.
XX
PN WO9913886-A1.
XX
PD 25-MAR-1999.
XX
PE 17-SEP-1998; 98WO-US19419.
XX
PR 09-JUN-1998; 98US-0093972.
XX
PR 17-SEP-1997; 97US-0059160.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 1999-229400/19.
XX
PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
PT vasoconstriction
XX
PS Disclosure; Page 37; 120pp; English.
XX
CC The specification describes antisense oligonucleotides (AA52869-55271)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of mRNAs corresponding to target genes, gene
CC initiation codons, genomic flanking regions, intron-exon borders, the
CC 5'-end, the 3'-end and the junction between coding and non-coding
CC regions and all segments of mRNAs encoding proteins associated with one
CC or more diseases, conditions or mixtures. The antisense oligonucleotides
CC may be derived from sequences AA52872-74. These multiple target
CC oligonucleotides (specifically AA52872-74) can be used for the
CC antisense treatment of diseases and conditions. Typical diseases and
CC conditions are those associated with impaired respiration and
CC inflammation, including lung diseases, pulmonary vasoconstriction,
CC inflammation, allergic rhinitis, acute asthma, allergic fibrosis,
CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
CC hepatic metastases, as well as all types of cancers which may metastasize
CC or have metastasized to the lungs, including breast and prostate cancer.
XX
SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;
Query Match 7.1%; Score 35.4; DB 20; Length 114955;
Best Local Similarity 32.9%; Pred. No. 3.9;
Matches 68; Conservative 23; Mismatches 116; Indels 0; Gaps 0;
Y 280 CCAGCGAGGCGCAACGCGGCGCTGATATGTTGGCCAGCGCCGACACGCAACCTGT 339
D 109573 CCACCCNNHNNNSGGGCGCCVAGNNHNNNSGGGCGCCVAGNNHNNNSGGGCGCCVAGG 109514

Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ds.
Homo sapiens.
WO200157182-A2.
09-AUG-2001.
17-JAN-2001; 2001WO-US01354.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0160528.
24-FEB-2000; 2000US-0184654.
02-MAR-2000; 2000US-0186350.
15-MAR-2000; 2000US-0189374.
17-MAR-2000; 2000US-0190076.
18-MAR-2000; 2000US-0198123.
19-MAR-2000; 2000US-0205513.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
18-AUG-2000; 2000US-0225759.
22-AUG-2000; 2000US-0226879.
22-AUG-2000; 2000US-0226888.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227182.
30-AUG-2000; 2000US-0228287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
08-SEP-2000; 2000US-0231242.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0231968.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0232397.
14-SEP-2000; 2000US-0232398.
14-SEP-2000; 2000US-0232399.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0232401.
14-SEP-2000; 2000US-0233063.
14-SEP-2000; 2000US-0233065.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
13-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239335.
13-OCT-2000; 2000US-0239337.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0248207.
17-NOV-2000; 2000US-0248208.
17-NOV-2000; 2000US-0248209.
17-NOV-2000; 2000US-0248210.
17-NOV-2000; 2000US-0248211.
17-NOV-2000; 2000US-0248212.
17-NOV-2000; 2000US-0248213.
17-NOV-2000; 2000US-0248214.
17-NOV-2000; 2000US-0248215.
17-NOV-2000; 2000US-0248216.
17-NOV-2000; 2000US-0248217.
17-NOV-2000; 2000US-0248218.
17-NOV-2000; 2000US-0248244.
17-NOV-2000; 2000US-0248245.
17-NOV-2000; 2000US-0248264.
17-NOV-2000; 2000US-0248265.
17-NOV-2000; 2000US-0248297.
17-NOV-2000; 2000US-0248299.
17-NOV-2000; 2000US-0248300.
01-DEC-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250191.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
06-DEC-2000; 2000US-0256719.
08-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251868.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
DR
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
PS
XX Disclosure; SEQ ID NO 38280; 3071pp + Sequence Listing; English.
CC
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
SQ
XX Sequence 16106 BP; 3275 A; 4801 C; 4361 G; 3669 T; 0 other;
Query Match 6.5%; Score 32.4; DB 22; Length 16106;
Best Local Similarity 56.6%; Pred. No. 17;
Matches 60; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
OY 29 TTGACGAGGTTCCAAAGTGTTCCTAAATTCGAAAGCGCCCTGCCCCCCTCCGCCCA 88
DB 1854 TTTCAAGAGTTCCAAATTTGAGTATTAATTTACGGCCAGCGCGCTCATGCTCTGTA 1795
OY 89 ATCTGCTTGCCTCGGGGGTGGGGGGTGGGGGCTCACCCTCTCAGG 134
DB 1794 ATCCCGACGACTTTGGGAGGCGCGGAGTGCAGATCAGATCAGG 1749
RESULT 12
AAK83469/C
ID AAK83469 standard; DNA; 16161 BP.
XX
XX AAK83469;
AC
XX
XX
DT 07-NOV-2001 (first entry)
XX
XX
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:38281.
XX
XX Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01354.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234297.
PR 25-SEP-2000; 2000US-0234974.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244826.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251858.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 38281; 3071bp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic

CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 16161 BP; 3307 A; 4825 C; 4360 G; 3669 T; 0 other;
 Query Match 6.5%; Score 32.4; DB 22; Length 16161;
 Best Local Similarity 56.6%; Pred. No. 17;
 Matches 60; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 29 TTGACAGGGTTGCAAGTTGTTCTTAATTCGAAGCGCCCTGCCCCCCTCCCA 88
 DB 1856 TTTCAGAGTTCCTCAATTTGAGTATGAATTTACGGCAGCGGGGTGCTCATGCCCTGTA 1797
 QY 89 ATCTGCTTGGTCGGGGGTGGGGGGTGGGGGGTCACTCTTCAGG 134
 DB 1796 ATCCAGCAGCTTTGGAGGCGGAGTGGGGGATCAGATCAGG 1751
 RESULT 13
 AAT94535
 ID AAT94535 standard; cDNA to mRNA; 1556 BP.
 AC AAT94535;
 XX 27-FEB-1998 (first entry)
 DT
 XX
 DE Maize cinnamoyl CoA reductase coding sequence.
 XX
 KW Alfaifa; cinnamoyl CoA reductase; eucaalyptus; transgenic plant; maize;
 KW ligin biosynthesis; ds.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT 195..1310
 FT CDS /tag= a
 FT /product= cinnamoyl CoA reductase
 XX
 PN W09712982-A1.
 XX
 PD 10-APR-1997.
 XX
 PE 03-OCT-1996; 96WO-FR01544.
 XX
 PR 03-OCT-1995; 95FR-0011623.
 XX
 PA (CNRS) CENT NAT RECH SCI.
 PA (INRG) INST NAT RECH AGRONOMICQUE.
 XX
 PI Beckert M, Boudet AM, Briat JF, Gamas P, Grima-Pettenati J;
 XX
 DR WPI; 1997-226225/20.
 DR P-PSDB; AAM36041.
 XX
 PT Transformation of plants to reduce or increase ligin content -
 PT using DNA encoding alfaifa or maize cinnamoyl CoA reductase or the
 PT corresponding anti-sense mRNA, e.g. to improve digestibility of
 PT fodder crops
 XX
 PS Claim 1; Page 38-40; 86pp; French.

XX This is the nucleotide sequence encoding maize cinnamoyl CoA reductase
CC (CCR). The sequence was isolated from a maize cDNA library using the
CC sequence of the eucalyptus CCR gene (AA094536). The CCR gene can be
CC used to transform plant cells to create transgenic plants in which
CC lignin biosynthesis is regulated so that lignin content is higher or
CC lower than in normal plants.
XX
SQ Sequence 1556 BP; 304 A; 500 C; 505 G; 247 T; 0 other:
Query Match 6.4%; Score 32.2; DB 18; Length 1556;
Best Local Similarity 45.2%; Pred. No. 8.7;
Matches 118; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 238 GGGGCTGGGGTGAACGAAAGGCTCCGAACTTTTTCAGCCAGGCGGAACGGG 297
DB 478 GCGGCGTCAAGGGTGCAGAGGCGCTTCCACACCGCTCCCGCTACCGACGACCGG 537
QY 298 GCGTCGTAATGATTGGCCAGGGGCGCATCGCAACCTGTCAATTCACGGGTCTCCGG 357
DB 538 AGCAAAATGTGGAGCGCGGGTGGGTTCACAGTGTGATCAACGCGGCGGAGG 597
QY 358 GTWGGAGGGGCGACGACCAAGCCCAACCCGCGGGAATCCGAGCAGTATATAGGGGCGCC 417
DB 598 CCGGACAGGTGCGGGGGTGGTTCACAGTGTGTCATCGGCGCGTGAACCATGACCCCA 657
QY 418 AGCTAGAGCCCAAGCAGACTGTGAATGCGACCTGTTCCAGAGAACTCATAGGTGCGAGA 477
DB 658 AGCGGCGGCGGACGCTGCTGTGCGACGAGTGTCTGCTGAGAGCAGCTCGATTCTCGAGA 717
QY 478 AGCGGCGGCGGTTCTGCTGAT 498
DB 718 AACCGAGAACTGTACTGCT 738
RESULT 14
AAH94477/C
ID AAH94477 standard; cDNA: 1177 BP.
XX
AC AAH94477;
XX
DT 05-OCT-2001 (first entry)
XX
DE Human foetal cDNA, SEQ ID NO: 1164.
XX
KW Human; foetal protein; cytosolic; immunosuppressive; immunostimulant;
KW neurotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
KW gene therapy; antisense therapy; cancer; immune disorder;
KW growth disorder; osteoporosis; thrombolytic disorder;
KW nervous system disorder; inflammation; expressed sequence tag; EST; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200155339-A2.
PN
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02723.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 15-SEP-2000; 2000US-0663870.
PR 06-NOV-2000; 2000US-0707351.
XX
XX (HYSE-) HYSEQ INC.
PI
XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
PI Liu C, Asundi V, Zhou P, Werhman T;
XX
XX WPI; 2001-465571/50.
DR P-PSDB; AAM06802.
XX
PT Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune

PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation -
XX
XX Example 3; Page 657-658; 715pp; English.
PS
XX The invention relates to novel foetal polypeptides encoded by
XX polynucleotides comprising one of 477 sequences fully defined in the
XX specification. The foetal polynucleotides and polypeptides are
XX useful in the treatment and diagnosis of diseases such as cancers,
XX immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
XX disorders, nervous system disorders and inflammation. The present
XX sequence is a full length cDNA which was assembled using expressed
XX sequence tags (ESTs) found to be expressed in human foetal tissue
XX cDNA libraries as seeds.
XX
SQ Sequence 1177 BP; 256 A; 291 C; 288 G; 336 T; 6 other:
Query Match 6.4%; Score 32; DB 22; Length 1177;
Best Local Similarity 52.2%; Pred. No. 9.1;
Matches 71; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 218 CCACGTCACTAGGCCAAGAGGGGCGTGGGTGAACGAAAGGCTCCGAACTTTT 277
DB 860 CGAGTTAACACGCGTGAACGAGAGGCTCAATCGACACGCTGTCACCAATGTGTTTC 801
QY 278 TTCAGGCGAGCGCGAAGCGGCGCTCGTATGATTTGGCCAGGCGCATCATCGGAACCT 337
DB 800 ACACGTGCTGCTGGGACATCGGCGGCTGGAGATCATGTGACCTCCAGAGAGACGTG 741
QY 338 GTCAATCACGGGTCTCT 353
DB 740 GGAAATATATGGTCTT 725

RESULT 15
AAF30438
ID AAF30438 standard; cDNA: 1623 BP.
XX
AC AAF30438;
XX
DT 29-MAY-2001 (first entry)
XX
DE Rice apoptosis inducing factor clone rds2c.pk002.g16 cDNA.
XX
KW Apoptosis inducing factor; AIF; monodehydroascorbate reductase;
KW male sterility; stress response; tissue culture; plant; rice; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FH CDS 51..1490
FT /*tag= a
XX
XX WO200119994-A2.
PN
XX
XX 22-MAR-2001.
PD
XX 11-SEP-2000; 2000WO-US24859.
PF
XX 13-SEP-1999; 99US-0153737.
PR
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA
XX Butler KH, Cahoon RE, Klein TM, Mazithulela G;
PI
XX WPI; 2001-244804/25.
DR P-PSDB; AAB20309.
XX
XX New polynucleotide encoding plant apoptosis inducing factor for
PT engineering male sterility, altering plant architecture, manipulating
PT stress response, and influencing growth of cells and tissues in culture
PT

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 18.733 Seconds
(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101a-10_COPY_1_500

Perfect score: 500

Sequence: 1 atccctccgcggcttcctca.....aagcgaagcattcattt 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	422	84.4	1500	1	US-08-704-398-1
2	422	84.4	1500	5	PCT-US95-0596-1
3	362.6	72.5	666	4	US-08-896-164-7
4	354.4	70.9	707	4	US-08-896-164-1
5	45.2	9.0	7218	1	US-08-232-463-14
6	32	6.4	6124	4	US-08-213-419-3
7	31.8	6.4	29793	4	US-09-302-812-38
8	31.8	6.4	29793	2	US-09-511-477-38
9	31.6	6.3	519	2	US-08-503-226B-36
10	31.6	6.3	519	3	US-08-721-458B-36
11	31.6	6.3	541	2	US-08-503-226B-35
12	31.6	6.3	541	3	US-08-721-458B-35
13	31.6	6.3	3588	1	US-08-197-792-32
14	31.6	6.3	3588	1	US-08-459-850-32
15	31.6	6.3	3588	1	US-08-459-214-32
16	31.6	6.3	8355	1	US-08-406-030A-23
17	31.2	6.2	897	5	PCT-US96-05320A-303
18	30.8	6.2	5852	1	US-07-867-106-2
19	30.4	6.1	3580	4	US-09-081-345-1
20	30	6.0	4403765	4	US-09-103-840A-2
21	30	6.0	4411529	4	US-09-103-840A-1
22	29.8	6.0	11464	4	US-08-991-840A-2
23	29.6	5.9	3457	1	US-08-295-882-1
24	29.4	5.9	2191	3	US-08-632-806A-6
25	29.4	5.9	2192	1	US-08-273-538A-6
26	29.4	5.9	3144	2	US-08-687-916-15
27	29.4	5.9	3144	4	US-09-138-614-15

28	29.4	5.9	11970	4	US-09-345-217-1	Sequence 1, Appl
29	29.4	5.9	152331	3	US-09-128-155-16	Sequence 16, Appl
30	29.2	5.8	627	4	US-09-385-982-502	Sequence 502, Appl
31	29.2	5.8	3396	4	US-08-974-549A-639	Sequence 639, Appl
32	29	5.8	1779	4	US-09-323-427-1	Sequence 1, Appl
33	29	5.8	1779	4	US-09-323-427-2	Sequence 2, Appl
34	29	5.8	2031	3	US-08-262-220-13	Sequence 13, Appl
35	29	5.8	2031	3	US-08-471-733-13	Sequence 13, Appl
36	29	5.8	2031	3	US-08-468-878-13	Sequence 13, Appl
37	29	5.8	2031	3	US-08-750-494-13	Sequence 13, Appl
38	29	5.8	2075	3	US-08-262-220-3	Sequence 3, Appl
39	29	5.8	2075	3	US-08-471-733-3	Sequence 3, Appl
40	29	5.8	2075	3	US-08-468-878-3	Sequence 3, Appl
41	29	5.8	2075	3	US-08-750-494-3	Sequence 3, Appl
42	28.8	5.8	1086	2	US-08-828-009-1	Sequence 1, Appl
43	28.8	5.8	1408	2	US-08-545-196B-11	Sequence 11, Appl
44	28.8	5.8	2763	4	US-09-030-335-3	Sequence 3, Appl
45	28.6	5.7	188	4	US-09-257-541-4	Sequence 4, Appl

ALIGNMENTS

Result 1
US-08-704-398-1
Sequence 1, Application US/08704398
Patent No. 5679525
GENERAL INFORMATION:
APPLICANT: Peterson, Michael G
TITLE OF INVENTION: EPSTEIN-BARR VIRUS TRANSCRIPTION FACTOR
TITLE OF INVENTION: BINDING ASSAY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,398
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/246,977
FILING DATE: 20-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET INFORMATION: A-59233/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1500 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1500
US-08-704-398-1
Query Match 84.4%; Score 422; DB 1; Length 1500;

Best Local Similarity 100.0%; Pred. No. 4.9e-124;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 79 ATGACACACGAGGAGGCTTGGCCGCGAGAGAGCCGCTGCGCATGCTCCATGCGCTGGG 138
    |||||||
Db 1 ATGACACACGAGGAGGCTTGGCCGCGAGAGAGCCGCTGCGCATGCTCCATGCGCTGGG 60
OY 139 AATTTGGTGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAATATTTTA 198
    |||||||
Db 61 AATTTGGTGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAATATTTTA 120
OY 199 AAGAGCGAGGAGGATCAACAGTACTTATCTCATGCAAAATGTCACAGAGTATAT 258
    |||||||
Db 121 AAGAGCGAGGAGGATCAACAGTACTTATCTCATGCAAAATGTCACAGAGTATAT 180
OY 259 GGAATATGAAAAAGTTTTTTTCCACCTCTTGTGTATCTTATGCGACGCGATGG 318
    |||||||
Db 181 GGAATATGAAAAAGTTTTTTTCCACCTCTTGTGTATCTTATGCGACGCGATGG 240
OY 319 AAGAAAAAAGAGACAAATGGAACGCGATGTTGTCTGAACAGAGTCTCAACCGTGT 378
    |||||||
Db 241 AAGAAAAAAGAGACAAATGGAACGCGATGTTGTCTGAACAGAGTCTCAACCGTGT 300
OY 379 GCATTTATTTGGATAGAAATAGTAGCAAGAAATGACAGAGTAACTGGAAGGAAG 438
    |||||||
Db 301 GCATTTATTTGGATAGAAATAGTAGCAAGAAATGACAGAGTAACTGGAAGGAAG 360
OY 439 AACTATTGACACGCAAAACATTTGTATATCTGACTGACAGACGGAAGCACTTCATT 498
    |||||||
Db 361 AACTATTGACACGCAAAACATTTGTATATCTGACTGACAGACGGAAGCACTTCATT 420
OY 499 TT 500
    ||
Db 421 TT 422

```

RESULT 2

PCT-US95-05966-1

; Sequence 1, Application PC/TUS9505966

; GENERAL INFORMATION:

; APPLICANT: TUDARIK, INC.

; TITLE OF INVENTION: EPSTEIN-BARR VIRUS TRANSCRIPTION FACTOR

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/05966

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/246,977

; FILING DATE: 20-MAY-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: FP-59233-PC/RAO

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 272799

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1500 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1500

; PCT-US95-05966-1

Query Match

Best Local Similarity 100.0%; Score 422; DB 5; Length 1500;

Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 79 ATGACACACGAGGAGGCTTGGCCGCGAGAGAGCCGCTGCGCATGCTCCATGCGCTGGG 138
    |||||||
Db 1 ATGACACACGAGGAGGCTTGGCCGCGAGAGAGCCGCTGCGCATGCTCCATGCGCTGGG 60
OY 139 AATTTGGTGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAATATTTTA 198
    |||||||
Db 61 AATTTGGTGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAATATTTTA 120
OY 199 AAGAGCGAGGAGGATCAACAGTACTTATCTCATGCAAAATGTCACAGAGTATAT 258
    |||||||
Db 121 AAGAGCGAGGAGGATCAACAGTACTTATCTCATGCAAAATGTCACAGAGTATAT 180
OY 259 GGAATATGAAAAAGTTTTTTTCCACCTCTTGTGTATCTTATGCGACGCGATGG 318
    |||||||
Db 181 GGAATATGAAAAAGTTTTTTTCCACCTCTTGTGTATCTTATGCGACGCGATGG 240
OY 319 AAGAAAAAAGAGACAAATGGAACGCGATGTTGTCTGAACAGAGTCTCAACCGTGT 378
    |||||||
Db 241 AAGAAAAAAGAGACAAATGGAACGCGATGTTGTCTGAACAGAGTCTCAACCGTGT 300
OY 379 GCATTTATTTGGATAGAAATAGTAGCAAGAAATGACAGAGTAACTGGAAGGAAG 438
    |||||||
Db 301 GCATTTATTTGGATAGAAATAGTAGCAAGAAATGACAGAGTAACTGGAAGGAAG 360
OY 439 AACTATTGACACGCAAAACATTTGTATATCTGACTGACAGACGGAAGCACTTCATT 498
    |||||||
Db 361 AACTATTGACACGCAAAACATTTGTATATCTGACTGACAGACGGAAGCACTTCATT 420
OY 499 TT 500
    ||
Db 421 TT 422

```

RESULT 3

US-08-896-164-7

; Sequence 7, Application US/08896164

; Patent No. 6218521

; GENERAL INFORMATION:

; APPLICANT: OBARA, Yuichi

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED

; WITH GASTRIC CANCER AND METHODS FOR

; DIAGNOSING AND TREATING GASTRIC CANCER

; NUMBER OF SEQUENCES: 87

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA.

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/896,164

; FILING DATE: July 17, 1997

; CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
 NAME: No. 6218521man D. Hanson
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 666 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-896-164-7

Query Match 72.5%; Score 362.6; DB 4; Length 666;
 Best Local Similarity 98.9%; Pred. No. 2.1e-105;
 Matches 365; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 132 GCCTGGGAATTTGGTGGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAA 191
 DB 31 GACAGGGAATTTGGTGGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAA 90
 QY 192 TTATTTAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGCACAGAA 251
 DB 91 TTATTTAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGCACAGAA 150
 QY 252 GTCATATGGAATGAAAAAGGTTTTTTTGGCCACCTCTTGATATATCTTATGCGCAG 311
 DB 151 GTCATATGGAATGAAAAAGGTTTTTTTGGCCACCTCTTGATATATCTTATGCGCAG 210
 QY 312 CGATGGAAGAAAAAAGAAATGAAAGGAGTGTGTTCTGAAAGAGCTCA 371
 DB 211 TCGATGGAAGAAAAAAGAAATGAAAGGAGTGTGTTCTGAAAGAGCTCA 270
 QY 372 ACCGTGCAATTTATTTGGATAGAAATAGTACCAAGAAATGACAGCTAAACTTGA 431
 DB 271 ACCGTGCAATTTATTTGGATAGAAATAGTACCAAGAAATGACAGCTAAACTTGA 330
 QY 432 AGGAAGAAGTATTTGACAGCCCAACATTTATATATCTGACTCAGACAGGAAAGCA 491
 DB 331 AGGAAGAAGTATTTGACAGCCCAACATTTATATATCTGACTCAGACAGGAAAGCA 390
 QY 492 CTCATTTT 500
 DB 391 CTCATGTT 399

RESULT 4
 US-08-896-164-1
 Sequence 1, Application US/08896164
 Patent No. 6218521
 GENERAL INFORMATION:
 APPLICANT: OBATA, YUICHI
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
 WITH GASTRIC CANCER AND METHODS FOR
 TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
 NUMBER OF SEQUENCES: 87
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felle & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WordPerfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/896,164
 FILING DATE: July 17, 1997

CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6218521man D. Hanson
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 707 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-896-164-1

Query Match 70.9%; Score 354.4; DB 4; Length 707;
 Best Local Similarity 97.0%; Pred. No. 8.6e-103;
 Matches 358; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 132 GCCTGGGAATTTGGTGGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAA 191
 DB 17 GACAGGGAATTTGGTGGAGCGGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAAA 76
 QY 192 TTATTTAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGCACAGAA 251
 DB 77 TTATTTAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGCACAGAA 136
 QY 252 GTCATATGGAATGAAAAAGGTTTTTTTGGCCACCTCTTGATATATCTTATGCGCAG 311
 DB 137 GTCATATGGAATGAAAAAGGTTTTTTTGGCCACCTCTTGATATATCTTATGCGCAG 196
 QY 312 CGATGGAAGAAAAAAGAAATGAAAGGAGTGTGTTCTGAAAGAGCTCA 371
 DB 197 TCGATGGAAGAAAAAAGAAATGAAAGGAGTGTGTTCTGAAAGAGCTCA 256
 QY 372 ACCGTGCAATTTATTTGGATAGAAATAGTACCAAGAAATGACAGCTAAACTTGA 431
 DB 257 ACCGTGCAATTTATTTGGATAGAAATAGTACCAAGAAATGACAGCTAAACTTGA 316
 QY 432 AGGAAGAAGTATTTGACAGCCCAACATTTATATATCTGACTCAGACAGGAAAGCA 491
 DB 317 AGGAAGAAGTATTTGACAGCCCAACATTTATATATCTGACTCAGACAGGAAAGCA 376
 QY 492 CTCATTTT 500
 DB 377 CTCATGTT 385

RESULT 5
 US-08-232-463-14/c
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 TITLE OF INVENTION: RECOMBINANT FOXP2 VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463

```

? RESULT 6
? US-08-213-419B-3/C
? Sequence 3, Application US/08213419B
? Patent No. 6333406
? GENERAL INFORMATION:
? APPLICANT: Inselburg, J. et al.
? TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
? FILE REFERENCE: JTI-00ZCNCP
? CURRENT APPLICATION NUMBER: US/08/213,419B
? CURRENT FILING DATE: 1994-03-17
? PRIOR APPLICATION NUMBERS: US 07/870,506
? PRIOR FILING DATE: 1993-04-17
? NUMBER OF SEQ ID NOS: 20
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 3
? LENGTH: 6124
? TYPE: DNA
? ORGANISM: Plasmodium falciparum

```

```

Best Local Similarity   59.3%; Pred. No. 11;
Matches    54; Conservative    0; Mismatches    37; Indels    0; Gaps    0

QY      296 TATATTCTTATGGCGAGCATGTGATGAATAAAAAAAGAACAATGAACCCGATGGTTCTT 355
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      14612 TATTTCGTGATGGAAGAAGATAGAGAAAATAAAACAGACTCTATATAAATGCATGTGATTC 14671

QY      356 CTGACAGAGAGCTCTCAACCCTGTGCGCTTTAT 386
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      14672 CTCGATTAATAATCCCTCTCTTTTTCACHTAT 14702

RESULT 8
US-09-511-477-38
Sequence 38, Application US/09511477
Patent No. 6357202
GENERAL INFORMATION:
APPLICANT: JACOBSON, MYRON K.
APPLICANT: JACOBSON, Elaine L.
```

APPLICANT: AME, Jean-Christophe
TITLE OF INVENTION: L1N, Winston
TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARC) EN
TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
FILE REFERENCE: N1AD 201
CURRENT APPLICATION NUMBER: US/09/511,477
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/302,812
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 38
LENGTH: 29793
TYPE: DNA
ORGANISM: Caenorhabditis elegans
FEATURE:
US-09-511-477-38

Query Match
Best Local Similarity 59.3%; Score 31.8; DB 4; Length 29793;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 236 TATATCTTATGCGCAGCGATGGAAGAAAAAGAACAAATGGAACCGATGCTTGT 355
DB 14612 TATTTCTGATGAGAGAGAGAGAGAAAAAGAGCTTAATTAATGATGATATC 14671
QY 356 CTGACAGAGCTCTCAACCGCTGTCATTAT 386
DB 14672 CTGATTAATAATCCCTCTTTTTCATTAAT 14702

RESULT 9
US-08-503-226B-36/c
Sequence 36, Application US/08503226B
Patent No. 5871945
GENERAL INFORMATION:
APPLICANT: Lockertle, Robert Owen, et al.
TITLE OF INVENTION: Calcineurin Inhibitory Compounds and
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/503,226B
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/404,731
FILING DATE: 15-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,227
FILING DATE: 23-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32861
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-6300
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-503-226B-36

Query Match
Best Local Similarity 49.4%; Score 31.6; DB 2; Length 519;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 128 CATCGCTGGAAATTTGGTAGCGGCTCCACTTAACGACTAGGAGGATATGC 187
DB 380 CAGGCCACGACCAACTGCTTACATTCCTGGCAGCTAGCTTCTTGGAAATCTGT 321
QY 188 GAATATTTTAAAGAGGAGGAGGATCAACAGTACTATTTTCATGCAAAAGTGCAC 247
DB 320 AAGTTGTTATGCTGCTACCTGTCAGCTCCAGCTGCTCTTTGGCAAACTTTCAA 261
QY 248 AGAAGCATATGCAAAATGAAAAAGTTTTCCTCCACCTCTTG 293
DB 260 GSAATTCCTACCCACCCAGAAAGTGTCTGTCAACCTGCATG 215

RESULT 10
US-08-721-458B-36/c
Sequence 36, Application US/08721458B
Patent No. 6107104
GENERAL INFORMATION:
APPLICANT: Lockertle, Robert Owen, et al.
TITLE OF INVENTION: Calcineurin Inhibitory Compounds
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,458B
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/404,731
FILING DATE: 15-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,227
FILING DATE: 23-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33276
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-6300
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-721-458B-36

Query Match
6.3%; Score 31.6; DB 3; Length 519;

Fri Oct 11 09:30:33 2002

us-09-489-101a-10_copy_1_500.rml

Page 8

Db 2409 GGGCCAGAAATAGCAAAAAAAAAAGAACCCCGTGGAGCCCGTGGCCAGAG 2468
QY 367 TCTCAACCGTGCGATTTATGGATAGCAAAATAGTACCAAGAAATGCAGCCTAAC 426
Db 2469 CCGGTCCCTTTTAACCAAGTAGGAGGGGAGTGAAGTGAATTTCTTGACA 2528
QY 427 TTGGAAGAGAAAGA 440
Db 2529 GTCCGAGCCAGAA 2542

RESULT 15
US-08-459-214-32
Sequence 32, Application US/08459214
Patent No. 5716810
GENERAL INFORMATION:
APPLICANT: Anthony J. Mason
INVENTOR: Peter H. Seedling
TITLE OF INVENTION: Nucleic Acid Encoding the Alpha or
TITLE OF INVENTION: Beta Chains of Inhibin and Method for Synthesizing Polypeptide
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: dattin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08459,214
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/197792
FILING DATE: 17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/958414
FILING DATE: 08-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/74207
FILING DATE: 12-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/215466
FILING DATE: 05-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/906729
FILING DATE: 31-DEC-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/827710
FILING DATE: 07-FEB-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/783910
FILING DATE: 03-OCT-1985
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 297P2D6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 3588 bases
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
US-08-459-214-32
Query Match 6.3%; Score 31.6; DB 1; Length 3588;
Best Local Similarity 52.2%; Pred. No. 3.9;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 307 GGCAGCGATGAGAGAAAAAAAAAGACAATGACGCGATGTTGTCGACAGAG 366
Db 2409 GGGCCAGAAATAGCAAAAAAAAAAGAACCCCGTGGAGCCCGTGGCCAGAG 2468
QY 367 TCTCAACCGTGCGATTTATGGATAGCAAAATAGTACCAAGAAATGCAGCCTAAC 426
Db 2469 CCGGTCCCTTTTAACCAAGTAGGAGGGGAGTGAAGTGAATTTCTTGACA 2528
QY 427 TTGGAAGAGAAAGA 440
Db 2529 GTCCGAGCCAGAA 2542

Search completed: October 10, 2002, 18:15:08
Job time: 58.7333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 695.733 Seconds
(without alignments)
9699.805 Million cell updates/sec

Title: US-09-489-101a-10_COPY_1_500

Sequence: 1 atccccccggttcttccca.....aagcgaagcactcatc 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estcm:*
5: em_estcv:*
6: em_estpl:*
7: em_estro:*
8: em_estro:*
9: em_estro:*
10: em_estro:*
11: em_estro:*
12: em_estro:*
13: em_estro:*
14: em_estro:*
15: em_estro:*
16: em_estro:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	478.8	95.4	595	10 BE257118	BE257118 601108855
2	477.2	95.4	662	10 BE253285	BE253285 601108755
3	475.6	95.1	716	10 BE269949	BE269949 601184864
4	475.2	95.0	785	10 BE1093196	BE1093196 602856157
5	474.6	94.9	690	10 BE513818	BE513818 601315659
6	454.2	90.8	658	10 BE269143	BE269143 601184493
7	423.8	84.8	524	10 BE269143	BE269143 601184493
8	383.4	76.7	473	10 BE609533	BE609533 323418 MA
9	382.2	76.4	575	10 BE1597820	BE1597820 603245301
10	382.2	76.4	748	10 BE1600065	BE1600065 603251452
11	382.2	76.4	699	10 BE1604083	BE1604083 603244010
12	378.2	75.6	463	10 BE226018	BE226018 u247c01.y
13	373	74.6	438	9 AA171575	AA171575 zp23a04.x
14	364.2	72.8	599	9 AL602969	AL602969 DKE2P686H
15	364.2	72.8	799	10 BG614689	BG614689 602642184
16	364.2	72.8	811	10 BG505015	BG505015 602551459
17	364.2	72.8	944	9 AL532999	AL532999 AL532999

18	364.2	72.8	947	9 AL520858	AL520858 AL520858
19	364.2	72.8	956	9 AL555008	AL555008 AL555008
20	362.6	72.5	636	9 AM239382	AM239382 kb39g08.y
21	362.6	72.5	846	9 AL556654	AL556654 AL556654
22	362.6	72.5	943	10 BE793102	BE793102 602254084
23	362.4	72.5	965	10 BG703001	BG703001 602664991
24	362.4	72.5	936	9 AL555798	AL555798 AL555798
25	362.4	72.5	1704	11 BC020780	BC020780 Homo sapi
26	361.8	72.4	749	9 AV710541	AV710541 AV710541
27	360.8	72.2	944	9 AL544062	AL544062 AL544062
28	360.4	72.1	683	9 AL580717	AL580717 AL580717
29	360.4	72.0	506	9 AL597064	AL597064 DKE2P313C
30	355	71.0	634	10 BE242424	BE242424 601876164
31	352.2	70.4	574	9 AL599175	AL599175 DKE2P3130
32	352.2	70.4	617	10 BE791029	BE791029 601586481
33	352	70.4	778	9 AL546160	AL546160 AL546160
34	351.4	70.3	839	10 BG623919	BG623919 602649076
35	351	70.2	855	9 AL516254	AL516254 AL516254
36	350	70.0	981	9 AL513760	AL513760 AL513760
37	350	69.6	984	10 BM472788	BM472788 ACENCO2T
38	347.8	69.6	947	10 BE783436	BE783436 601471756
39	347.4	69.5	810	10 B1837899	B1837899 603083718
40	347.2	69.4	569	10 BG828970	BG828970 602752978
41	347	69.4	714	10 BG433877	BG433877 602497615
42	345.8	69.2	747	10 B1836305	B1836305 603083385
43	344.8	69.0	740	10 BE537860	BE537860 601067322
44	339	67.8	769	10 B1827542	B1827542 603074237
45	338.8	67.8	646	10 BE884239	BE884239 601505704

ALIGNMENTS

RESULT 1
LOCUS BE257118 595 bp mRNA linear EST 13-JUL-2000
DEFINITION BE257118 601108855F1 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:3349813 5',
RNA sequence.
ACCESSION BE257118
VERSION BE257118.1 GI:9127588
KEYWORDS EST.
SOURCE human
ORGANISM human
Human sapiens
Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS NIH_MGC_16
TITLES NIH_MGC_16
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: genops-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LINC44 row: m column: 14
High quality sequence stop: 595.
Location/Qualifiers
1. 595
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3349813"
/clone_lib="NIH_MGC_16"
/tissue="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pORF7; Site: 1; XhoI; Site: 2;
ECORI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GCCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit

FEATURES

source
1. 595
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3349813"
/clone_lib="NIH_MGC_16"
/tissue="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pORF7; Site: 1; XhoI; Site: 2;
ECORI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GCCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit

BASE COUNT 185 a 124 c 158 g 128 t
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH-MGC library.

Query Match 95.8%; Score 478.8; DB 10; Length 595;
 Best Local Similarity 99.6%; Pred No. 1.6e-108;
 Matches 480; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CAGCTCCACGACGAGGCTCCCAAGGCGCTCAAAACCGGATACCGGAGCGCTCCG 78
 |||||||
 DB 57 CAGCTCCACGACGAGGCTCCCAAGGCGCTCAAAACCGGATACCGGAGCGCTCCG 116
 |||||||

QY 79 ATGACACACACGAGGAGGCTGCGCGGAGGAGCGCGCTGGCATGCTCATGCGCTGG 138
 |||||||
 DB 117 ATGACACACACGAGGAGGCTGCGCGGAGGAGCGCGCTGGCATGCTCATGCGCTGG 176
 |||||||

QY 139 AAATTGGTGAAGGCTCCACCTAAACGACTAGAGGAAAGCTATGCGAATATTATTA 198
 |||||||
 DB 177 AAATTGGTGAAGGCTCCACCTAAACGACTAGAGGAAAGCTATGCGAATATTATTA 236
 |||||||

QY 199 AAGAGGAGGAGGATCAACAGTACTATCTTCATGCAAAAGTTGACAGAGATCAT 258
 |||||||
 DB 237 AAGAGGAGGAGGATCAACAGTACTATCTTCATGCAAAAGTTGACAGAGATCAT 296
 |||||||

QY 259 GGAATATGAAAAAGGTTTTTTTGGCCACGCTCTGTGTATATCTTATGCGAGCGATG 318
 |||||||
 DB 297 GGAATATGAAAAAGGTTTTTTTGGCCACGCTCTGTGTATATCTTATGCGAGCGATG 356
 |||||||

QY 319 AAGAAAAAAGAAAGAAATGGAAGCGATGTTGTGTGAAAGAGTCTCAACGCTGT 378
 |||||||
 DB 357 AAGAAAAAAGAAAGAAATGGAAGCGATGTTGTGTGAAAGAGTCTCAACGCTGT 416
 |||||||

QY 379 GCATTATTTGGGATAGAAATATGACCAAGAAATGACAGAGCTAAACTTGGAAGGAAG 438
 |||||||
 DB 417 GCATTATTTGGGATAGAAATATGACCAAGAAATGACAGAGCTAAACTTGGAAGGAAG 476
 |||||||

QY 439 AACTATTGACACGCAAAACATTTGTATATATCTGACTGACAGCAAGGAAAGCACTTCAT 498
 |||||||
 DB 477 AACTATTGACACGCAAAACATTTGTATATATCTGACTGACAGCAAGGAAAGCACTTCAT 536
 |||||||

QY 499 TT 500
 ||
 DB 537 TT 538
 ||

RESULT 2
 LOCUS BE253285 662 bp mRNA linear EST 13-JUL-2000
 DEFINITION 60110875F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3349812 5',
 mRNA sequence.
 ACCESSION BE253285
 VERSION BE253285.1 GI:9123350
 KEYWORDS EST.
 SOURCE human;
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 662) NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3349812 5',
 NIH-MGC http://mgc.ncl.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rps@post.rockefeller.edu
 Tissue: Protoplasts
 Procedure: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNase Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
 High quality sequence stop: 341.
 Location/Qualifiers

source
 1. 662
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3349812"
 /clone_id="NIH_MGC_16"
 /issue="type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site: 1; XhoI; Site: 2;
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCGACGAG(6). Library constructed by Ling Hong
 in the Laboratory of Gerald W. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH-MGC library."

BASE COUNT 207 a 143 c 169 g 143 t

Query Match 95.4%; Score 477.2; DB 10; Length 662;
 Best Local Similarity 99.4%; Pred No. 4.1e-108;
 Matches 479; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 CAGTCTCAGTACGAGGCTCCCAAGGCGCTCAAAACCGGATACCGGAGCGCTCCG 78
 |||||||
 DB 56 CAGTCTCAGTACGAGGCTCCCAAGGCGCTCAAAACCGGATACCGGAGCGCTCCG 115
 |||||||

QY 79 ATGACACACACGAGGAGGCTGCGCGGAGGAGCGCGCTGGCATGCTCATGCGCTGG 138
 |||||||
 DB 116 ATGACACACACGAGGAGGCTGCGCGGAGGAGCGCGCTGGCATGCTCATGCGCTGG 175
 |||||||

QY 139 AAATTGGTGAAGGCTCCACCTAAACGACTAGAGGAAAGCTATGCGAATATTATTA 198
 |||||||
 DB 176 AAATTGGTGAAGGCTCCACCTAAACGACTAGAGGAAAGCTATGCGAATATTATTA 235
 |||||||

QY 199 AAGAGGAGGAGGATCAACAGTACTATCTTCATGCAAAAGTTGACAGAGATCAT 258
 |||||||
 DB 236 AAGAGGAGGAGGATCAACAGTACTATCTTCATGCAAAAGTTGACAGAGATCAT 295
 |||||||

QY 259 GGAATATGAAAAAGGTTTTTTTGGCCACGCTCTGTGTATATCTTATGCGAGCGATG 318
 |||||||
 DB 296 GGAATATGAAAAAGGTTTTTTTGGCCACGCTCTGTGTATATCTTATGCGAGCGATG 355
 |||||||

QY 319 AAGAAAAAAGAAAGAAATGGAAGCGATGTTGTGTGAAAGAGTCTCAACGCTGT 378
 |||||||
 DB 356 AAGAAAAAAGAAAGAAATGGAAGCGATGTTGTGTGAAAGAGTCTCAACGCTGT 415
 |||||||

QY 379 GCATTATTTGGGATAGAAATATGACCAAGAAATGACAGAGCTAAACTTGGAAGGAAG 438
 |||||||
 DB 416 GCATTATTTGGGATAGAAATATGACCAAGAAATGACAGAGCTAAACTTGGAAGGAAG 475
 |||||||

QY 439 AACTATTGACACGCAAAACATTTGTATATATCTGACTGACAGCAAGGAAAGCACTTCAT 498
 |||||||
 DB 476 AACTATTGACACGCAAAACATTTGTATATATCTGACTGACAGCAAGGAAAGCACTTCAT 535
 |||||||

QY 499 TT 500
 ||
 DB 536 TT 537
 ||

RESULT 3
 LOCUS BE269949 716 bp mRNA linear EST 13-JUL-2000
 DEFINITION 60118486F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3542655 5',
 mRNA sequence.
 ACCESSION BE269949
 VERSION BE269949.1 GI:9143578
 KEYWORDS EST.
 SOURCE human;
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 716) NIH_MGC http://mgc.ncl.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMML at: Image.liml.gov
Plate: L10M234 row: P column: 16
High quality sequence stop: 042.

FEATURES
Source
1. 718 sm="Homo sapiens"
/organism="Homo sapiens"
/db_name="taxon:9606"
/clone_id="IMAGE:3542655"
/tissue_type="BaurKit Lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph. Vector: pORF7. Site: 1: XhoI; site: 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 218 a 155 c 193 g 150 t
ORIGIN

Query Match 95.1%; Score 475.6; DB 10; Length 716;
Best Local Similarity 99.2%; Pred. No. 1e-107;
Matches 476; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 19 CAGCTCCACGACGCTCCCTCAAGCGGCTCTAAACCGGATACCGGAGCGCTCCG 78
DB 86 CAGCTCCACGACGCTCCCTCAAGCGGCTCTAAACCGGATACCGGAGCGCTCCG 145
OY 79 ATGGACACACGAGGAGGCTCCGCGGAGAGCGGCTCCGATGCTCCATGCGCTGG 138
DB 146 ATGGACACACGAGGAGGCTCCGCGGAGAGCGGCTCCGATGCTCCATGCGCTGG 205
OY 139 AATTTGTGAGGCGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAATTTATTA 198
DB 206 AATTTGTGAGGCGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAATTTATTA 265
OY 199 AAGAGCGAGGAGGATCAACAGTACTTATTTCTGCAAAAGTTGACAGAGATCAT 258
DB 266 AAGAGCGAGGAGGATCAACAGTACTTATTTCTGCAAAAGTTGACAGAGATCAT 325
OY 259 GGAATATGAAAAAGTTTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 318
DB 326 GGAATATGAAAAAGTTTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 385
OY 319 AAGAAAAAAGAAAAAGTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 378
DB 386 AAGAAAAAAGAAAAAGTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 445
OY 379 GATTTTGTGATGAGAAATAGTGAACAAGAGAGAGTAACTTGAAGAAG 438
DB 446 GATTTTGTGATGAGAAATAGTGAACAAGAGAGAGTAACTTGAAGAAG 505
OY 439 AACTTATGACAGCAGCAAAACATTTATATATCTGACTACAGCAAGAGAGATCAT 498
DB 506 AACTTATGACAGCAGCAAAACATTTATATCTGACTACAGCAAGAGAGATCAT 565
OY 499 TT 500
DB 566 TT 567

BI093196
LOCUS BI093196 785 bp mRNA linear EST 20-JUN-2001
DEFINITION 602856137F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4997295 5',
ACCESSION BI093196
VERSION BI093196.1 GI:14511526
KEYWORDS EST.
SOURCE Human
ORGANISM Human sapiens
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
TITLES 1 (bases 1 to 785)
AUTHORS Mammalia; Eutheria; Primates; Chordata; Granulata; Vertebrata; Euteleostomi;
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMML at: http://image.liml.gov
Plate: L10M1024 row: b column: 16
High quality sequence stop: 755.

FEATURES
Source
1. 785
/organism="Homo sapiens"
/db_name="taxon:9606"
/clone_id="IMAGE:4997295"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-Sport6; Site: 1: NotI; site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 Kb. Library prepared by Life Technologies."

BASE COUNT 243 a 168 c 198 g 173 t 3 others
ORIGIN

Query Match 95.0%; Score 475.2; DB 10; Length 785;
Best Local Similarity 99.2%; Pred. No. 1.3e-107;
Matches 477; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 19 CAGCTCCACGACGCTCCCTCAAGCGGCTCTAAACCGGATACCGGAGCGCTCCG 78
DB 78 CAGCTCCACGACGCTCCCTCAAGCGGCTCTAAACCGGATACCGGAGCGCTCCG 137
OY 79 ATGGACACACGAGGAGGCTCCGCGGAGAGCGGCTCCGATGCTCCATGCGCTGG 138
DB 138 ATGGACACACGAGGAGGCTCCGCGGAGAGCGGCTCCGATGCTCCATGCGCTGG 197
OY 139 AATTTGTGAGGCGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAATTTATTA 198
DB 198 AATTTGTGAGGCGCTCCACCTAAACGACTTACTAGGAGAGCTATGCGAATTTATTA 257
OY 199 AAGAGCGAGGAGGATCAACAGTACTTATTTCTGCAAAAGTTGACAGAGATCAT 258
DB 258 AAGAGCGAGGAGGATCAACAGTACTTATTTCTGCAAAAGTTGACAGAGATCAT 317
OY 259 GGAATATGAAAAAGTTTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 318
DB 318 GGAATATGAAAAAGTTTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 377
OY 319 AAGAAAAAAGAAAAAGTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 378
DB 378 AAGAAAAAAGAAAAAGTTTGGCCACCTCTTGATATCTTATGCGACGCGATGG 437
OY 379 GATTTTGTGATGAGAAATAGTGAACAAGAGAGTAACTTGAAGAAG 438
DB 438 GATTTTGTGATGAGAAATAGTGAACAAGAGAGTAACTTGAAGAAG 497

RESULT 4

```

QY 439 AACTATTCACAGCCAAACATTTATATATCTGACACAGAACGCAACTTCATT 498
DB 498 AACTATTCACAGCCAAACATTTATATATCTGACACAGAACGCAACTTCATT 557
QY 499 T 499
DB 558 T 558

RESULT 5
BE513818
LOCUS BE513818
DEFINITION 690 bp mRNA linear EST 07-AUG-2000
MENA sequence.
mRNA sequence.
BE513818
EST.
EST. GI:9721030

ACCESSION BE513818
VERSION BE513818.1
KEYWORDS GI:9721030

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 690)
NHI-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM238 row: 3 column: 16
High quality sequence stop: 656.
Location/Qualifiers
1. 690
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3634191"
/clone_lib="NHI-MGC-8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 211 a 145 c 104 g 145 t
ORIGIN

Query Match 94.9%; Score 474.6; DB 10; Length 690;
Best Local Similarity 99.2%; Pred. No. 1.8e-107;
Matches 477; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 19 CAGCTCCACGACGCTCCCTCAAAAGCGGCTCTAAACCCGATTAACCGAGCGCTCCC 78
DB 87 CAGCTCCACGACGCTCCCTCAAAAGCGGCTCTAAACCCGATTAACCGAGCGCTCCC 146
QY 79 ATGACACACGAGAGCGCTGCGCGGAGAGAGCGCGCTCCATGCTCCATGCGCTGGG 138
DB 147 ATGACACACGAGAGCGCTGCGCGGAGAGAGCGCGCTCCATGCTCCATGCGCTGGG 206
QY 139 AAATTTGTGACGCGCTCCCTCAAAAGCGCTCTAAACCCGATTAACCGAGCGCTCCC 198
DB 207 AAATTTGTGACGCGCTCCCTCAAAAGCGCTCTAAACCCGATTAACCGAGCGCTCCC 266
QY 139 AAATTTGTGACGCGCTCCCTCAAAAGCGCTCTAAACCCGATTAACCGAGCGCTCCC 258
DB 267 AAATTTGTGACGCGCTCCCTCAAAAGCGCTCTAAACCCGATTAACCGAGCGCTCCC 326

```

```

QY 259 GGAATGAAAAAGGTTTTTGGCCACCTCTTGTATATCTTATGGCAGCGATGG 318
DB 337 GGAATGAAAAAGGTTTTTGGCCACCTCTTGTATATCTTATGGCAGCGATGG 366
QY 319 AAGAAAAAAGAAAGAAATGAGACGCGATGTTGTTCTGAACAAGCTCTCAACCGTGT 378
DB 387 AAGAAAAAAGAAAGAAATGAGACGCGATGTTGTTCTGAACAAGCTCTCAACCGTGT 446
QY 379 GCATTTATGAGATGAGAAATGATGACCAAGAAATGACGAGCTAACTTGAAGAGAA 438
DB 447 GCATTTATGAGATGAGAAATGATGACCAAGAAATGACGAGCTAACTTGAAGAGAA 506
QY 439 AACTATTCACAGCCAAACATTTATATATCTGACACAGAACGCAACTTCATT 498
DB 507 AACTATTCACAGCCAAACATTTATATATCTGACACAGAACGCAACTTCATT 566
QY 499 T 499
DB 567 T 567

RESULT 6
BE269143
LOCUS BE269143
DEFINITION 658 bp mRNA linear EST 13-JUL-2000
MENA sequence.
mRNA sequence.
BE269143
EST.
EST. GI:9142758

ACCESSION BE269143
VERSION BE269143.1
KEYWORDS GI:9142758

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 658)
NHI-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM233 row: 3 column: 24
High quality sequence stop: 583.
Location/Qualifiers
1. 658
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="354215"
/clone_lib="NHI-MGC-8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 195 a 136 c 182 g 145 t
ORIGIN

Query Match 90.8%; Score 454.2; DB 10; Length 658;
Best Local Similarity 99.0%; Pred. No. 2.1e-102;
Matches 478; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 19 CAGCTCCACGACGCTCCCTCAAAAGCGGCTCTAAACCCGATTAACCGAGCGCTCCC 78

```


Db 87 CAGTCTCCAGCACTGCTCCCTCAAGCGGCTCTTAAACCCGGATACCGGAGCGCTCC 146
 Oy 79 ATGACACACAGGAGGCTTCCCGCGAGAGACCCCTCCGCAATGCTCATGCGCTGG 138
 Db 147 ATGACACACAGGAGGCTTCCCGCGAGAGACCCCTCCGCAATGCTCATGCGCTGG 206
 Oy 139 AAATTTGGTACGCGCTCCACCTAAAGCACTACTAGAGAAAGTATGCAATATATTA 198
 Db 207 AAATTTGGTACGCGCTCCACCTAAAGCACTACTAGAGAAAGTATGCAATATATTA 266
 Oy 199 AAAGAGGAGGAGGATCAACAGTACTATCTTATGCAAGAAAGTATGCAAGTATAT 258
 Db 267 AAAGAGGAGGAGGATCAACAGTACTATCTTATGCAAGAAAGTATGCAAGTATAT 326
 Oy 259 GGAATTTGAAAAAGGTTTTTTTGGCCACCTCTGTATATCTTATGCAAGGAGGATGG 318
 Db 327 GGAATTTGAAAAAGGTTTTTTTGGCCACCTCTGTATATCTTATGCAAGGAGGATGG 386
 Oy 319 AAGAAAAAAGAAAAAGCAATATGCAAGGAGGATGGTGTGTAAGCAAGAGTCTGT 378
 Db 387 AAGAAAAAAGAAAAAGCAATATGCAAGGAGGATGGTGTGTAAGCAAGAGTCTGT 446
 Oy 379 GCATTTATTTTGGGATGAGAAATAGTACCAAGAAATGCAAGGATGGTGTGTAAG 437
 Db 447 GCATTTATTTTGGGATGAGAAATAGTACCAAGAAATGCAAGGATGGTGTGTAAG 506
 Oy 438 GAATCTATGCAAGCAATCTGTATATCTGATGCAAGCAAGGATGGTGTGTAAG 497
 Db 507 GAATCTATGCAAGCAATCTGTATATCTGATGCAAGCAAGGATGGTGTGTAAG 565
 Oy 498 TTT 500
 Db 566 GTT 568

RESULT 7
 LOCUS BE269331 524 bp mRNA linear EST 13-JUL-2000
 DEFINITION 601185204P1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3542833 5',
 mRNA sequence.
 ACCESSION BE269331
 VERSION BE269331.1 GI:9142950
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 524)
 NIH-MGC <http://mgi.ncl.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMU at: image.lmu.gov
 Plate: L10K235 row: h column: 02
 High quality sequence stop: 524.
 Location/Qualifiers
 1 524
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3542833"
 /clone_lib="NIH_MGC_8"
 /tissue_type="Burkitt Lymphoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lymph. Vector: pOT87; Site: 1. XhoI; Site: 2.
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average

Insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald W. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

Query Match 84.8%; Score 423.8; DB 10; Length 524;
 Best Local Similarity 99.3%; Pred. No. 7.3e-95;
 Matches 436; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

BASE COUNT 158 a 113 c 147 g 106 t
 ORIGIN
 19 CAGTCTCCAGCACTGCTCCCTCAAGCGGCTCTTAAACCCGGATACCGGAGCGCTCC 78
 Db 86 CAGTCTCCAGCACTGCTCCCTCAAGCGGCTCTTAAACCCGGATACCGGAGCGCTCC 145
 Oy 79 ATGACACACAGGAGGCTTCCCGCGAGAGACCCCTCCGCAATGCTCATGCGCTGG 138
 Db 146 ATGACACACAGGAGGCTTCCCGCGAGAGACCCCTCCGCAATGCTCATGCGCTGG 205
 Oy 139 AAATTTGGTACGCGCTCCACCTAAAGCACTACTAGAGAAAGTATGCAATATATTA 198
 Db 206 AAATTTGGTACGCGCTCCACCTAAAGCACTACTAGAGAAAGTATGCAATATATTA 265
 Oy 199 AAAGAGGAGGAGGATCAACAGTACTATCTTATGCAAGAAAGTATGCAAGTATAT 258
 Db 266 AAAGAGGAGGAGGATCAACAGTACTATCTTATGCAAGAAAGTATGCAAGTATAT 325
 Oy 259 GGAATTTGAAAAAGGTTTTTTTGGCCACCTCTGTATATCTTATGCAAGGAGGATGG 318
 Db 326 GGAATTTGAAAAAGGTTTTTTTGGCCACCTCTGTATATCTTATGCAAGGAGGATGG 385
 Oy 319 AAGAAAAAAGAAAAAGCAATATGCAAGGAGGATGGTGTGTAAGCAAGAGTCTGT 378
 Db 386 AAGAAAAAAGAAAAAGCAATATGCAAGGAGGATGGTGTGTAAGCAAGAGTCTGT 445
 Oy 379 GCATTTATTTTGGGATGAGAAATAGTACCAAGAAATGCAAGGATGGTGTGTAAG 437
 Db 446 GCATTTATTTTGGGATGAGAAATAGTACCAAGAAATGCAAGGATGGTGTGTAAG 505
 Oy 438 GAATCTATGCAAGCAATCTGTATATCTGATGCAAGCAAGGATGGTGTGTAAG 497
 Db 506 GAATCTATGCAAGCAATCTGTATATCTGATGCAAGCAAGGATGGTGTGTAAG 565

RESULT 8
 LOCUS BG609533 473 bp mRNA linear EST 17-APR-2001
 DEFINITION 323418 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BG609533
 VERSION BG609533.1 GI:13659512
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 REFERENCE 1 (bases 1 to 473)
 Fahnenkrug, S.C., Fekking, B.A., Rohrer, G.A., Smith, P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grose, W.M., Bennett, G.A., Laegreid, W.W.
 and Keele, J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 Contact: Smith TPJ
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@meat.mars.usda.gov
 Single Pass sequencing. Passes called and alt-trimmed with phred
 v0.960904.e. Vector identified by cross-match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT

full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
Library.
BASE COUNT 252 a 146 c 194 g 207 t
ORIGIN

Query Match 72.8%; Score 364.2; DB 10; Length 799;
Best Local Similarity 99.2%; Pred. No. 4.4e-80;
Matches 366; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
OY 132 GCGTGGGAATTTGGTGGAGCGGCTCCACCTTAAGCACTTACTAGGGAGCTATGCGAA 191
    |||||||
DB 38 GACAGGGAATTTGGTGGAGCGGCTCCACCTTAAGCACTTACTAGGGAGCTATGCGAA 97
    |||||||
OY 192 TTATTTAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGCACAGAA 251
    |||||||
DB 98 TTATTTAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGCACAGAA 157
    |||||||
OY 252 GTCATATGGAATGAAAAAGTTTTTTTGGCCACCTCCTGTGTATATCTTATGGGCG 311
    |||||||
DB 158 GTCATATGGAATGAAAAAGTTTTTTTGGCCACCTCCTGTGTATATCTTATGGGCG 217
    |||||||
OY 312 CCGATGGAAGAAAAAACAATGGAAGCGGATGTTGTTGTAACAAGACTCTCA 371
    |||||||
DB 218 CGGATGGAAGAAAAAACAATGGAAGCGGATGTTGTTGTAACAAGACTCTCA 277
    |||||||
OY 372 ACCGTGTGCAATTTATGGGATAGAAAAATAGTACCAAGAAATGCACAGCTAACTTGA 431
    |||||||
DB 278 ACCGTGTGCAATTTATGGGATAGAAAAATAGTACCAAGAAATGCACAGCTAACTTGA 337
    |||||||
OY 432 AGGAAAGACTATGTCACAGCAAAACATGTATATCTGACTGACAGCAAGCAAGCA 491
    |||||||
DB 338 AGGAAAGACTATGTCACAGCAAAACATGTATATCTGACTGACAGCAAGCAAGCA 397
    |||||||
OY 492 CTTCATTTT 500
    |||||||
DB 398 CTTCATGTT 406
```

Search completed: October 10, 2002, 20:27:56
Job time : 702.733 secs

Result No.	Score	Query Length	DB	ID	Description
1	500	100.0	1580	20	AAK39671 Renal cancer assoc
2	500	100.0	1580	22	AD01117 Human small cell 1
3	432.8	86.6	1581	13	AAO20595 Jk recombinant sldbrn
4	422	84.4	1500	17	AAI06947 C-promoter binding
5	364.2	72.8	5658	21	AAPI8323 lung cancer associ
6	362.6	70.5	666	20	AAK40082 Gastric cancer ass
7	354.4	70.9	707	20	AAK40076 Gastric cancer ass
8	348.8	69.8	1336	22	AAK27790 Human prostate car
9	327.2	65.4	373	22	AAH2783 Nucleotide sequen

C	10	320.8	64.2	159.9	20	AAH33641	Breast cancer asso
	11	319.6	63.9	47.3	22	AAH33426	Human colon cancer
	12	312.6	62.5	66.5	22	AAFE2796	Human prostate can
	13	283.8	56.6	76.8	22	AAFE2795	Human prostate can
	14	281.8	56.4	62.5	22	AAFE2792	Human prostate can
	15	238.8	47.8	81.5	22	AAFE2791	Human prostate can
	16	213.8	42.8	39.1	22	AAH55516	Human breast tumor
	17	181.2	36.2	22.9	21	AAAC00786	Human secreted pro
	18	171.4	34.3	35.8	21	AAAC79290	Human lung tumor
	19	171.4	34.3	37.8	21	AAFD23366	Human lung tumor
	20	160.8	32.2	74.1	22	AAFE2793	Human prostate can
	21	127.4	25.5	29.3	22	AAH06262	Drosophila melanos
	22	115.2	23.0	25.1	19	AAH1526	Human biallelic pol
	23	115.2	23.0	25.1	19	AAH1527	Human biallelic pol
	24	115.2	23.0	26.2	20	AAH87430	Human single nucle
	25	115.2	23.0	26.2	20	AAH87431	Human single nucle
	26	111	22.2	58.39	23	AAH06268	Drosophila melanos
	27	67	13.4	30.3	21	AAAC08043	Human secreted pro
	28	65.2	13.0	11.37	22	AAH85004	Human transcritp
	29	65.2	13.0	23.70	21	AACT7028	Human ORFX ORF258
	30	47.4	9.5	81.9	22	AAFE2649	Human gastric can
C	31	41.2	8.2	27.6	23	AAST7162	DNA encoding novel
	32	41.2	8.2	36.62	23	AAST7162	DNA encoding novel
	33	41.2	8.2	36.62	23	AAAG59571	DNA encoding novel
	34	35.6	7.1	81.9	21	AAZ35392	Maize starch bran
	35	35	7.0	32.367	19	AAH35620	Human SHOX (short
	36	34.6	6.9	89.64	24	AAH33534	Human immune syst
	37	34.6	6.9	89.64	24	AAH561276	Human immune syst
	38	34.4	6.9	115.6	20	AAZ15396	Human gene regulat
	39	34.4	6.9	314.8	20	AAH34436	Human gene express
C	40	34.4	6.9	345.0	22	AAH34587	S. epidermidis gen
	41	34.4	6.9	376.9	22	AAH34587	S. epidermidis gen
C	42	34	6.8	185.98	24	AAH32387	Human immune syst
C	43	33.8	6.8	269.38	22	AAH82620	Human HBM gene reg
C	44	33.8	6.8	106.47	20	AAH561396	Human gene regulat
	45	33.8	6.8	136.11	21	AAFE21288	Human low adenosti

ALIGNMENTS

RESULT 1	ID	AAK39671	standard; DNA; 1580 BP.
XX	AC	AAK39671;	
XX	DT	02-JUL-1999	(first entry)
XX	DE	Renal cancer associated gene.	
XX	KW	Cancer associated antigen; diagnosis; research; treatment; human; breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer; prostate cancer; ss.	
XX	OS	Homo sapiens.	
XX	PN	W09904265-A2.	
XX	PD	28-JAN-1999.	
XX	PF	15-JUL-1998;	98W0-US14679.
XX	PR	22-JUN-1998;	98US-0102322.
XX	PR	17-JUL-1997;	97US-0886164.
XX	PR	10-OCT-1997;	97US-0061592.
XX	PR	10-OCT-1997;	97US-0061765.
XX	PR	10-OCT-1997;	97US-0948705.
XX	PR	11-OCT-1997;	97GB-0021697.
XX	PA	(LUDW-) LUDWIG INST CANCER RES.	
XX	CH	Chen Y, Gout I, Gure A, O'Hare M, Ohta Y, Old LJ;	

PI Pfrendschuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX WPI: 1999-132448/11.
XX
XX New isolated cancer associated nucleic acids and polypeptides -
PT Isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
PS Claim 67: Page 481-482; 787pp: English.
XX
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule, and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
XX Sequence 1580 BP; 481 A; 352 C; 359 G; 388 T; 0 other;
XX
Query Match 100.0%; Score 500; DB 20; Length 1580;
Best Local Similarity 100.0%; Pred. No. 2, 3e-146;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
1 ATCCCGCCGGTTTCTCAGTCTCCAGTACGTCCTCAAGCGGCTCTAAACCCGG 60
DB 1 ATCCCGCCGGTTTCTCAGTCTCCAGTACGTCCTCAAGCGGCTCTAAACCCGG 60
XX
61 ATACCGGAGGCGCTCCCATGACGACGACGAGGCTTGGCCCGGAGAGCCGCTGG 120
DB 61 ATACCGGAGGCGCTCCCATGACGACGACGAGGCTTGGCCCGGAGAGCCGCTGG 120
XX
121 CATGCTCATCGCTGCGAAATTTGGTGAGCGGCTTCACCTTAACGACTTACAGGAA 180
DB 121 CATGCTCATCGCTGCGAAATTTGGTGAGCGGCTTCACCTTAACGACTTACAGGAA 180
XX
181 GCTATGCGAAATTTATTAAGAGGAGGATCAACAGTACTTATCTTCATGCAAAA 240
DB 181 GCTATGCGAAATTTATTAAGAGGAGGATCAACAGTACTTATCTTCATGCAAAA 240
XX
241 GTTGACAGAGATCATATGAAATGAAATGTTTGTGGCCACCTCTTGCTATAT 300
DB 241 GTTGACAGAGATCATATGAAATGAAATGTTTGTGGCCACCTCTTGCTATAT 300
XX
301 CTATGAGGCGAGGATGAGAAAGAAAGAAAGAAAGAAAGAGATGATGTTGTTGAA 360
DB 301 CTATGAGGCGAGGATGAGAAAGAAAGAAAGAAAGAAAGAGATGATGTTGTTGAA 360
XX
361 CAGAGATGCAACCGTGGCATTTTGGGATAGAAATGTACCAAGAAATGACAGAG 420
DB 361 CAGAGATGCAACCGTGGCATTTTGGGATAGAAATGTACCAAGAAATGACAGAG 420
XX
421 CTAAACTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 480
DB 421 CTAAACTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 480
XX
481 AAGCGAAGACACTTATTTT 500
DB 481 AAGCGAAGACACTTATTTT 500
XX
RESULT 2
AAD1117
ID AAD1117 standard: DNA; 1580 BP.

AC AAD1117;
XX
XX 24-SEP-2001 (first entry)
DE Human small cell lung cancer associated gene, RBP-1.
XX
XX Human; small cell lung cancer; therapy: hCAAP; nucleic acid; NA;
KW melanoma; cancer; colon; breast; head; neck; transitional cancer;
KW leiomyosarcoma; J-kappa recombination signal binding protein; RBP-1;
XX synovial sarcoma; cytostatic; ds.
XX
OS Homo sapiens.
XX
XX WO200153349-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001MO-US02015.
XX
XX 21-JAN-2000; 2000US-0489101.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (SLOK) SLOAN KETTERING INST CANCER RES.
XX (CORR) CORNELL RES FOUND INC.
XX
XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
XX WPI: 2001-457597/49.
XX
XX Isolated polypeptide, used to treat or prognose a disorder
PT characterized by expression of a hCAAP e.g. cancer, is encoded by an
PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
XX
XX Claim 57; Page 97; 152pp: English.
XX
XX The invention relates to nucleic acids and encoded polypeptides which
XX are cancer associated antigens expressed in patients afflicted with
XX small cell lung cancer. The molecules provided by the invention can be
XX used in the diagnosis, monitoring, research or treatment of conditions
XX characterised by the expression of one or more cancer associated
XX antigens. The polypeptide is used to treat a disorder characterised by
XX expression of a hCAAP, and determine regression, progression or onset
XX of a condition characterised by expression of an abnormal amount of a
XX protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
XX are small and non-small cell lung cancer, melanoma, colon, breast, head
XX and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
XX The present sequence is a DNA encoding human J-kappa recombination
XX signal binding protein (RBP-1). This small cell lung cancer associated
XX gene is designated as NY-SCLC-8.
XX
XX Sequence 1580 BP; 481 A; 352 C; 359 G; 388 T; 0 other;
XX
Query Match 100.0%; Score 500; DB 22; Length 1580;
Best Local Similarity 100.0%; Pred. No. 2, 3e-146;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
1 ATCCCGCCGGTTTCTCAGTCTCCAGTACGTCCTCAAGCGGCTCTAAACCCGG 60
DB 1 ATCCCGCCGGTTTCTCAGTCTCCAGTACGTCCTCAAGCGGCTCTAAACCCGG 60
XX
61 ATACCGGAGGCGCTCCCATGACGACGACGAGGCTTGGCCCGGAGAGCCGCTGG 120
DB 61 ATACCGGAGGCGCTCCCATGACGACGACGAGGCTTGGCCCGGAGAGCCGCTGG 120
XX
121 CATGCTCATCGCTGCGAAATTTGGTGAGCGGCTTCACCTTAACGACTTACAGGAA 180
DB 121 CATGCTCATCGCTGCGAAATTTGGTGAGCGGCTTCACCTTAACGACTTACAGGAA 180
XX
181 GCTATGCGAAATTTATTAAGAGGAGGATCAACAGTACTTATCTTCATGCAAAA 240
DB 181 GCTATGCGAAATTTATTAAGAGGAGGATCAACAGTACTTATCTTCATGCAAAA 240
XX
241 GTTGACAGAGATCATATGAAATGAAATGTTTGTGGCCACCTCTTGCTATAT 300
DB 241 GTTGACAGAGATCATATGAAATGAAATGTTTGTGGCCACCTCTTGCTATAT 300

selectively bind to a transcription factor (TF) (such as Epstein-Barr virus EBNA2), and a test compound. The sample is incubated under conditions that allow the protein and the TF to bind. The presence or absence of binding of CBP1 and TF is then detected. Absence of binding indicates that the test compound disrupts CBP1-TF dependent gene expression. This method is also modified to include a labeled nucleic acid sequence containing the sequence GGA in the incubation mixture. When the nucleic acid is used, the presence or absence of the binding of the CBP1-TF transcription complex to the nucleic acid sequence is detected. The test compounds identified are potentially useful for treating diseases associated with expression of a gene mediated by a CBP1-TF transcription complex. These include viral, bacterial or fungal infections, metabolic or genetic diseases and cell growth/regulatory dysfunction (such as neoplasms, inflammation or hypersensitivity). In humans animals or plants.

Sequence 1500 BP; 465 A; 319 C; 346 G; 370 T; 0 other;

Query Match 84.4%; Score 422; DB 17; Length 1500;
Best Local Similarity 100.0%; Pred. No. 6.5e-122;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 ATGGACACACGAGGAGGCTTGGCCGAGAGAGCCGCTGCGATGCTCCATCCCTGG 138
DB 1 ATGGACACACGAGGAGGCTTGGCCGAGAGAGCCGCTGCGATGCTCCATCCCTGG 60
QY 139 AATTTGGTGGAGGCGCTCCACCTAAACGACTTCTAGGAGAGCTTGGGAATTTATTA 198
DB 61 AATTTGGTGGAGGCGCTCCACCTAAACGACTTCTAGGAGAGCTTGGGAATTTATTA 120
QY 199 AAGAGCGAGGAGGATCAACAGTACTTATTTCTATGCAAAAGTTGACAGAGCATAT 258
DB 121 AAGAGCGAGGAGGATCAACAGTACTTATTTCTATGCAAAAGTTGACAGAGCATAT 180
QY 259 GGAATGAAAGAGTTTGGCCACCTGCTGTATATTTTGGGCGAGCATGG 318
DB 181 GGAATGAAAGAGTTTGGCCACCTGCTGTATATTTTGGGCGAGCATGG 240
QY 319 AAGAAAGAGAGAGTAAAGTGAAGAGGAGTGGTCTGCAAGAGGCTCAACCGCT 378
DB 241 AAGAAAGAGAGTAAAGTGAAGAGGAGTGGTCTGCAAGAGGCTCAACCGCT 300
QY 379 GCATTTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 438
DB 301 GCATTTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
QY 439 AACTATTGACAGCCAAACATTTGATATCTGACTGACAGAGGAAAGCACTTCATT 498
DB 361 AACTATTGACAGCCAAACATTTGATATCTGACTGACAGAGGAAAGCACTTCATT 420
QY 499 TT 500
DB 421 TT 422

RESULT 5
AAF18323
ID AAF18323 standard; DNA; 5668 BP.
XX
XX AAF18323;
XX
XX 14-MAR-2001 (first entry)
XX
XX Lung cancer associated polynucleotide sequence SEQ ID 342.
XX
XX Human: lung cancer associated protein; neuroprotective; cytosolic;
XX cardioactive; immunomodulatory; muscular active; vulnery;
XX gastrointestinal; nephrotoxic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease; ds.
XX
XX Homo sapiens.
XX

PN WO20005180-A2.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-0505918.
XX 12-MAR-1999; 99US-0124270.
XX (HUMAN) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX Ruben SM;
XX WPI; 2000-587514/55.
XX P-PDSB; AAB58447.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as lung cancer -
XX Claim 1; Page 798-800; 1425pp; English.

Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
associated proteins represented in AAB58106 - AAB58548. Lung cancer
associated proteins and polynucleotide sequences, their agonists, and
antagonists may have neuroprotective; cytosolic; cardioactive;
immunomodulatory; muscular active general; vulnery; gastrointestinal
activity; nephrotoxic; antiinfective; gynecological; or antibacterial
activity. The invention also includes antibodies specific for the
protein or polynucleotide sequences. The lung cancer associated
polynucleotide sequences may be used for detection of lung cancer,
chromosome identification, as chromosome markers, and for numerous other
diagnostic or research purposes. The proteins may be used to treat
disorders such as neural, immune, muscular, reproductive
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
disorders. The proteins may also be used in the treatment of wounds and
infectious diseases. Polynucleotide sequences AAF18425 - AAF18435 and
peptide AAB58549 are used in the course of the invention for the
identification and characterization of the polynucleotide and protein
sequences.

Sequence 5668 BP; 1647 A; 1028 C; 1113 G; 1879 T; 1 other;

Query Match 72.8%; Score 364.2; DB 21; Length 5668;
Best Local Similarity 99.2%; Pred. No. 1.7e-103;
Matches 366; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 132 GCGTGGAAATTTGGTGAAGGCGCTCCACCTAAAGCACTTACTAGGAGAGCTTGGGAA 191
DB 183 GACAGGAAATTTGGTGAAGGCGCTCCACCTAAAGCACTTACTAGGAGAGCTTGGGAA 242
QY 192 TTATTTAAAGAGGAGGAGGATCAACAGTACTTATTTCTATGCAAAAGTTGACAGAA 251
DB 243 TTATTTAAAGAGGAGGAGGATCAACAGTACTTATTTCTATGCAAAAGTTGACAGAA 302
QY 252 GTCATATGAAATGAAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 311
DB 303 GTCATATGAAATGAAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 362
QY 312 CGATGGAAG 371
DB 363 CGATGGAAG 422
QY 372 ACCGTGCACTTTTGGGATGAGAAATGAGCAACAGAAATGAGCAACAGCACTTGGGA 431
DB 423 ACCGTGCACTTTTGGGATGAGAAATGAGCAACAGAAATGAGCAACAGCACTTGGGA 482
QY 432 AGGAAAGAACTATTGACAGCAGCAAAACATTTGATATCTGACTGACAGAGAGAGAG 491
DB 483 AGGAAAGAACTATTGACAGCAGCAAAACATTTGATATCTGACTGACAGAGAGAGAG 542
QY 492 CTTGATTTT 500
111111 11

DB 543 CTCATGTT 551

RESULT 6
AA40082
ID AA40082 standard; DNA; 666 BP.
XX
XX
AC AA40082;
XX
XX 02-JUL-1999 (first entry)
DE
XX Gastric cancer associated gene.
XX
XX Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
XX Homo sapiens.
OS
XX WO904265-A2.
PN
XX 28-JAN-1999.
PD
XX
XX 15-JUL-1998; 98WO-US14679.
PF
XX
XX 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ,
PI Pfeundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
XX WPI: 1999-132448/11.
XX
XX New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
XX
PS Claim 67; Page 691; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
XX
SQ Sequence 666 BP; 214 A; 120 C; 155 G; 173 T; 4 other;

Query Match 72.5%; Score 362.6; DB 20; Length 666;
Best Local Similarity 98.9%; Pred. No. 1.8e-103;
Matches 365; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 132 GCGTGGAAATTTGGTGGAGGCGCTCCACCTTAAGCACTTACTGAGGAAGCTATGGGAAA 191
DB 31 GACAGGGAATTTGGTGGAGGCGCTCCACCTTAAGCACTTACTGAGGAAGCTATGGGAAA 90
OY 192 TATTTTAAAGGAGGAGGAGTCAACACTACTTATTTCTTATTCAGAAAATTCACAGAA 251
|||||

DB 91 TTTATTTAAAGAGCGAGGAGCAACAGACTTATTTCTTCATGCAAAAATTGCACAGAA 150

OY 252 GTCATATGAAATGATAAAGGTTTATTTGGCCACCTCTCTGTGTATCTTATCTTANGGGAG 311
DB 151 GTCATATGAAATGATAAAGGTTTATTTGGCCACCTCTCTGTGTATCTTATCTTANGGGAG 210
OY 312 CGGATGGAATAAATAAAGCAATGAGAGCGATGCTGTGTCTGACAGAGACTCA 371
DB 211 TGGATGGAATAAATAAAGCAATGAGAGCGATGCTGTGTCTGACAGAGACTCA 270
OY 372 ACCGTGTGCAATTTATTTGGATAGCAATAGTGCACCAAGAAATGACACTTAACTTGA 431
DB 271 ACCGTGTGCAATTTATTTGGATAGCAATAGTGCACCAAGAAATGACACTTAACTTGA 330
OY 432 AGCAAGACACTTATTTGCACAGCCCAAACTTATTTATTTCTGACTCAGACAGCAAGCA 491
DB 331 AGCAAGACACTTATTTGCACAGCCCAAACTTATTTATTTCTGACTCAGACAGCAAGCA 390

OY 492 CTCATTTT 500
DB 391 CTCATGTT 399

RESULT 7
AA40076
ID AA40076 standard; DNA; 707 BP.
XX
XX
AC AA40076;
XX
XX 02-JUL-1999 (first entry)
DE
XX Gastric cancer associated gene.
XX
XX Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
XX Homo sapiens.
OS
XX WO904265-A2.
PN
XX 28-JAN-1999.
PD
XX
XX 15-JUL-1998; 98WO-US14679.
PF
XX
XX 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ,
PI Pfeundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
XX WPI: 1999-132448/11.
XX
XX New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
XX
PS Claim 67; Page 689-689; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a

XX 18-JAN-2000; 2000GB-0000993.
 PR (UNO-) UNIV NOTTINGHAM TRENT.
 PA
 XX Rees RC, LI G, Mian S;
 PI
 XX WPI: 2001-476121/51.
 DR
 XX Use of cancer-associated genes and their products in detecting,
 PT monitoring, treating or preventing cancer, specifically prostate
 PT cancer, and in developing DNA-based vaccines that promote anti-tumor
 PR responses -
 PS
 XX Claim 1; Page 18; 43pp; English.
 CC AAH42781-AAH42846 represent the nucleotide sequences of cancer
 CC associated genes, identified using SEREX (Serological Identification of
 CC Antigens by Recombinant Expression Cloning). The genes are overexpressed
 CC in prostate cancer, and some are overexpressed in other cancers such as
 CC stomach cancer and oesophageal cancer. The nucleic acids are useful for
 CC detecting cancer and monitoring cancer, particularly prostate cancer. They are
 CC particularly useful in the treatment or prevention of cancer. In
 CC producing DNA-based vaccines against prostate cancer or that promote
 CC anti-tumor immune responses, and to raise antibodies. The expression of
 CC genes and detection of their protein products and/or peptides may be
 CC used to monitor disease progression during therapy, or as a prognostic
 CC indicator of initial disease status of the patient.
 XX
 SO Sequence 373 BP; 128 A; 59 C; 93 G; 93 T; 0 other;
 Query Match 65.4%; Score 327.2; DB 22; Length 373;
 Best Local Similarity 98.8%; Pred. No. 1,7e-97;
 Matches 340; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 132 GCGTGGGAAATTTGGTGGAGCGGCTCCACCTTAACGACTTACTAGGAGAGCTATGCGA 191
 DB 30 GACAGGGAATTTGGTGGAGCGGCTCCACCTTAACGACTTACTAGGAGAGCTATGCGA 89
 QY 192 TTAATTTAAAGACGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGACAGAA 251
 DB 90 TTAATTTAAAGACGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGACAGAA 149
 QY 252 GTCAATATGGAATGAAAGGTTTGGCCACCTCTGTGTATATCTTATGCGAG 311
 DB 150 GTCAATATGGAATGAAAGGTTTGGCCACCTCTGTGTATATCTTATGCGAG 209
 QY 312 CGGATGGAAGAAAAAAGAACAAATGGAACGGATGTTCTTCAACAGACTCA 371
 DB 210 TCGATGGAAGAAAAAAGAACAAATGGAACGGATGTTCTTCAACAGACTCA 269
 QY 372 ACCGTGCACTTTATGGATAGGAAATAGTACCAAGAAATGAGAGCTTAACCTTGA 431
 DB 270 ACCGTGCACTTTATGGATAGGAAATAGTACCAAGAAATGAGAGCTTAACCTTGA 329
 QY 432 AGGAAGAACTATGCAAGCCCAAAACATTTATATATCTGACT 475
 DB 330 AGGAAG-ACTATGCAAGCCCAAAACATTTATATCTGACT 372
 RESULT 10
 AAX39641
 ID AAX39641 standard; DNA; 1599 BP.
 XX
 AC AAX39641;
 XX
 DT 02-JUL-1999 (first entry)
 XX
 DE Breast cancer associated gene.
 XX
 KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer; ss.

XX Homo sapiens.
 OS
 XX MO9904265-A2.
 PN
 XX 28-JAN-1999.
 PD
 XX 15-JUL-1998; 98MO-US14679.
 PF
 XX 22-JUN-1998; 98US-0102322.
 PR 17-JUL-1997; 97US-0896164.
 PR 10-OCT-1997; 97US-0061599.
 PR 10-OCT-1997; 97US-0061765.
 PR 10-OCT-1997; 97US-0948705.
 PR 11-OCT-1997; 97GB-0021697.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Chen Y, Gout I, Gure A, O'Hare M, Ohta Y, Old LJ;
 PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
 PI Tureci O;
 XX
 XX WPI: 1999-132448/11.
 DR
 XX New isolated cancer associated nucleic acids and polypeptides -
 PT isolated using sera from cancer patients, used to develop products
 PT for the diagnosis, monitoring or treatment of cancers
 PR
 XX
 PS Claim 67; Page 401-402; 787pp; English.
 CC The invention relates to a method for diagnosing a disorder characterised
 CC by expression of a human cancer associated antigen precursor coded for by
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 CC biological sample, isolated from a subject with an agent that specifically
 CC binds to the NAM, an expression product or a fragment of an expression
 CC product complexed with an RNA molecule; and (b) determining the
 CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer.
 XX
 SO Sequence 1599 BP; 498 A; 317 C; 358 G; 426 T; 0 other;
 Query Match 64.2%; Score 320.8; DB 20; Length 1599;
 Best Local Similarity 92.6%; Pred. No. 3,7e-90;
 Matches 337; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 137 GGAATTTGGTGGAGCGGCTCCACCTTAACGACTTACTAGGAGAGCTATGCGA 196
 DB 137 GGAATTTGGTGGAGCGGCTCCACCTTAACGACTTACTAGGAGAGCTATGCGA 196
 QY 197 TAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGACAGAACTCA 256
 DB 197 TAAAGAGCGAGGGGATCAACAGTACTTATCTTCATGCAAAAGTTGACAGAACTCA 256
 QY 257 ATGGAATGAAAGGTTTGGCCACCTCTGTGTATATCTTATGCGAGCGGAT 316
 DB 257 ATGGAATGAAAGGTTTGGCCACCTCTGTGTATATCTTATGCGAGCGGAT 316
 QY 317 GGAAGAAAAAAGAACAAATGGAACGGATGTTCTTCAACAGACTCA 376
 DB 317 GGAAGAAAAAAGAACAAATGGAACGGATGTTCTTCAACAGACTCA 376
 QY 377 GTGCACTTTATGGATAGGAAATAGTACCAAGAAATGAGAGCTTAACCTTGA 436
 DB 377 GTGCACTTTATGGATAGGAAATAGTACCAAGAAATGAGAGCTTAACCTTGA 436
 QY 437 AGAATATGCAAGCCCAAAACATTTATATATCTGACTCAAGAGAGGAGCACTTCA 496
 DB 437 AGAATATGCAAGCCCAAAACATTTATATATCTGACTCAAGAGAGGAGCACTTCA 496

DB 437 AGAAGTATGACAGCAACAAACATGTGACATATGATTCAGACAGCAAGCAAGCACTTCA 436
 QY 437 TTTT 500
 DB 437 TTTT 500
 DB 437 TTTT 500
 RESULT 11
 AAH33426
 ID AAH33426 standard; cDNA; 473 BP.
 XX
 AC AAH33426;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:482.
 XX
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma, ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US6524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 XX WPI: 2001-235357/24.
 DR P-PSDB; AAG73995.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 1: Page 2571; 9803pp; English.
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated PS,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 XX Sequence 473 BP; 142 A; 76 C; 136 G; 114 T; 5 other:
 Query Match 63.9%; Score 319.6; DB 22; Length 473;
 Best Local Similarity 97.2%; Pred. No. 4,7e-90;
 Matches 341; Conservative 4; Mismatches 4; Indels 2; Gaps 2;
 QY 132 GCCGCGGAATTTGGTGAGCGGCTCCACCTTAAGAGATTCTCTGAGGAGAGCTATGCGAAA 191
 DB 91 GACAGGGAATTTGGTGAGCGGCTCCACCTTAAGAGATTCTCTGAGGAGAGCTATGCGAAA 150

QY 192 TTTTAAAGAGCGGAGGATCAACAGATGCTTATTTCTTCATGCAAAAAGTTGCACAGAA 251
 DB 151 TTTTAAAGAGCGGAGGATCAACAGATGCTTATTTCTTCATGCAAAAAGTTGCACAGAA 210
 QY 252 GTCATATGGAATGAAAAAGGTTTTTTTGGCCACCTCTTGTATATCTATGCGGAG 311
 DB 211 GTCATATGGAATGAAAAAGGTTTTTTTGGCCACCTCTTGTATATCTATGCGGAG 270
 QY 312 CGGATGGAAGAAAAAAGAAAAAAGATGAGCGGATGTTGTTGTGACAAAGAGTCGA 371
 DB 271 YGATGGAAGAAAAAAGAAAAAAGATGAGCGGATGTTGTTGTGACAAAGAGTCGA 330
 QY 372 ACCGTGCAATTTATGGAATGAGAAATAGTACCAAGAAATGACAGCTAACTT-CG 430
 DB 331 ACCGTGCAATTTATGGAATGAGAAATAGTACCAAGAAATGACAGCTAACTTGGG 390
 QY 431 AAGGAAGAACTATTGACAGCGCAAAACATGTTATATCTGACTGACGA 481
 DB 391 AAGGAAGAACTATTGACAGCGC-AAACMTGTATATCTGCTGACGA 440
 RESULT 12
 AAF22796
 ID AAF22796 standard; cDNA; 665 BP.
 XX
 AC AAF22796;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen nucleotide sequence SEQ ID:375.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine, ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US14749.
 XX
 PR 28-MAY-1999; 99US-0136526.
 PR 10-SEP-1999; 99US-0153454.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PI Odata Y;
 XX
 XX WPI: 2001-025274/03.
 DR
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 XX cancer -
 XX
 PS Claim 50: Page 397; 799pp; English.
 XX
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 XX represent nucleotide sequences encoding human breast, gastric and
 XX prostate cancer associated antigen precursors (CAAP) respectively.
 XX AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 XX represent human breast, gastric and prostate CAAP protein sequence
 XX respectively. CAAPs have cytostatic activity and can be used in the
 XX production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 XX acids or anti-CAAP antibodies are useful for diagnosing and treating a
 XX condition characterized by expression of an abnormal amount of a protein,
 XX e.g. cancer.
 SQ
 Query Match 62.5%; Score 312.6; DB 22; Length 665;
 Best Local Similarity 96.6%; Pred. No. 8,9e-88;

Matches 340; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 137 GGAATTTGTTAGGGGCTCCTAACGACTTACTAGGAGGATGAGCAATTTAT 196
 |||||
 Db 211 GGAATTTGTTAGGGGCTCCTAACGACTTACTAGGAGGATGAGCAATTTAT 270
 QY 197 TAAAAGACGAGGAGTCAACACTTATTTCTTCATGCAAAAGTTCACAGAACTCAT 256
 |||||
 Db 271 TAAAAGACGAGGAGTCAACACTTATTTCTTCATGCAAAAGTTCACAGAACTCAT 330
 QY 257 ATGGAATGAAAAAGTTTTTTTGGCCACCTCTGTGTATATCTTATGAGGACGGAT 316
 |||||
 Db 331 ATGGAATGAAAAAGTTTTTTTGGCCACCTCTGTGTATATCTTATGAGGACGGAT 390
 QY 317 GGAAGAAAAAAGAACCAATGAGACGCGATGTTCTGTGAACAAGATCTCAACCGT 376
 |||||
 Db 391 GGAAG-AAAAAAGAACCAATGAGACGCGATGTTCTGTGAACAAGATCTCAACCGT 449
 QY 377 GTGCATTTATTTGGATAGCAATATGACCAAGAAATGCAAGCTAACTGGAAGAA 436
 |||||
 Db 450 GTGCATTTATTTGGATAGCAATATGACCAAGAAATGCAAGCTAACTGGAAGAA 509
 QY 437 AGAATATTGACAGCCAAACATTGTATATCTGTGACTGACAGCAAGCGAAA 488
 |||||
 Db 510 AGAATATTGACACA-CCAAACATTGTATATCTGTGACTGACAGCAAGCGAAA 560

RESULT 13

AAF22795
 ID AAF22795 standard; cDNA: 768 BP.

XX AAF22795;

DT 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen nucleotide sequence SRO ID:374.

DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytosolic; cancer vaccine; ss.

XX Homo sapiens.

PN WO200073801-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US14749.

PR 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

PI MPI: 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.

XX cancer -

XX Claim 50; Page 396-397; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP), respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterized by expression of an abnormal amount of a protein,

CC e.g. cancer.

XX Sequence 768 BP; 254 A; 139 C; 169 G; 204 T; 2 other;

SO Query Match 56.6%; Score 282.8; DB 22; Length 768;

Best Local Similarity 99.3%; Pred. No. 2, 1e-78;

Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 215 AATGCACTTATTTCTTCATGCAAAAGTTCACAGAGTCAATATGCAATGAAAAAGCT 274
 |||||
 Db 1 AATGCACTTATTTCTTCATGCAAAAGTTCACAGAGTCAATATGCAATGAAAAAGCT 60
 QY 275 TTTTGGCCACCTCTGTGTATATCTTATGAGGACGCGATGGAAGAAAAAAGAAC 334
 |||||
 Db 61 TTTTGGCCACCTCTGTGTATATCTTATGAGGACGCGATGGAAGAAAAAAGAAC 120
 QY 335 AATGGAACGCGATGTTCTGTGAACAAGATCTCAACCGTGTGATTTATGAGATAG 394
 |||||
 Db 121 AATGGAACGCGATGTTCTGTGAACAAGATCTCAACCGTGTGATTTATGAGATAG 180
 QY 395 GAATATGAGACCAAGAAATGCAAGCTAACTTGAAGGAAAGACTTTTCACAGCA 454
 |||||
 Db 181 GAATATGAGACCAAGAAATGCAAGCTAACTTGAAGGAAAGACTTTTCACAGCA 240
 QY 455 AATGATTTATATCTGACTGACAGCAAGGAAAGCACTTATTT 500
 |||||
 Db 241 AATGATTTATATCTGACTGACAGCAAGGAAAGCACTTATTT 286

RESULT 14

AAF22792
 ID AAF22792 standard; cDNA: 625 BP.

XX AAF22792;

DT 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen nucleotide sequence SRO ID:371.

DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytosolic; cancer vaccine; ss.

XX Homo sapiens.

PN WO200073801-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US14749.

PR 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

PI MPI: 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.

XX cancer -

XX Claim 50; Page 395-396; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP), respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterized by expression of an abnormal amount of a protein,
e.g. cancer.
XX

SQ Sequence 625 BP; 205 A; 103 C; 142 G; 164 T; 11 other;

Query Match 56.4%; Score 281.8; DB 22; Length 625;
Best Local Similarity 91.8%; Pred No. 3.9e-78;
Matches 315; Conservative 0; Mismatches 25; Indels 3; Gaps 2;

QY 137 GGAATTTGGTGAAGCGCTCCACTTAAGACTTACAGAGAGCTATGCGAATATTAT 196
DB 208 GGAATTTGGTGAAGCGCTCCACTTAAGACTTACAGAGAGCTATGCGAATATTAT 267
QY 197 TAAAGAGCGAGGGGATCAACAGTACTTATCTTCAATGCAAAAGTGCACAGAGCAT 256
DB 268 TAAAGAGCGAGGGGATCAACAGTACTTATCTTCAATGCAAAAGTGCACAGAGCAT 327
QY 257 ATGGAATGAAAAAGTTTTTTTGGCCACCTCTGTGTATATCTATGAGGAGCGGAT 316
DB 328 ATGGAATGAAAAAGTTTTTTTGGCCACCTCTGTGTATATCTTATGAGGAGGAT 387
QY 317 GGAAGAAAAAAGAACAAATGGAACGCGATGTTGTTGACAGAGA--GTCTCAAC 374
DB 388 GGAAGAAAAAAGAACAAATGGAACGCGATGTTGTTGACAGAGA--GTCTCAAC 447
QY 375 GTGTGCTTTTATTTGGATAGGAATAGTGACCAAGAAATGACAGAGCTTAACTTGAAG 434
DB 448 GGTGTGCTTTTATTTGGATAGGAATAGTGACCAAGAAATGACAGAGCTTAACTTGAAG 506
QY 435 AAGAACTATTGTCACAGCCAAACATTTGATATATCTGACTCA 477
DB 507 AAGAACTATTGTCACAGCCAAACATTTGATATATCTGACTCA 549

RESULT 15

AAF22791
ID AAF22791 standard; cDNA; 815 BP.

XX AAF22791;

XX AC 26-MAR-2001 (first entry)

XX DT Human prostate cancer associated antigen nucleotide sequence SEQ ID:370.

XX DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

XX KM cancer associated antigen; cytostatic; cancer vaccine; ss.

XX OS Homo sapiens.

XX PN W0200073801-A2.

XX PD 07-DEC-2000.

XX PF 26-MAY-2000; 2000MO-US14749.

XX PR 28-MAY-1999; 99US-0136526.

XX PR 10-SEP-1999; 99US-0153454.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Obata Y;

XX DR WPI; 2001-025274/03.

XX PT Nucleic acids encoding breast, gastric and prostate cancer associated
XX PT antigen precursors, useful for diagnosing and treating a condition
XX PT characterized by expression of an abnormal amount of a protein, e.g.
XX PT cancer -

XX PS Claim 50: Page 395; 799pp; English.

XX CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterized by expression of an abnormal amount of a protein,
e.g. cancer.
XX

SQ Sequence 815 BP; 261 A; 149 C; 179 G; 218 T; 8 other;

Query Match 47.8%; Score 238.8; DB 22; Length 815;
Best Local Similarity 98.8%; Pred. No. 1.4e-64;
Matches 251; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 247 CAGAGTCATATGGAATGAAAAAGTTTTTTGGCCACCTCTGTGTATATCTATG 306
DB 1 CAGAGTCATATGGAATGAAAAAGTTTTTTGGCCACCTCTGTGTATATCTATG 60
QY 307 GCGAGCGATGGAAGAAAAAAGAACAAATGGAACGCGATGTTGTTCTGACAGAG 366
DB 61 GCGAGCGATGGAAG--AAAAAAGAACAAATGGAACGCGATGTTGTTCTGACAGAG 119
QY 367 TCTCAACCGTGTGCAATTTATTTGGATAGGAATAGTGACCAAGAAATGACAGCTAAAC 426
DB 120 TCTCAACCGTGTGCAATTTATTTGGATAGGAATAGTGACCAAGAAATGACAGCTAAAC 179
QY 427 TTGGAAGGAAGAACTATGTCACAGCCAAACATTTGATATCTGACTCAGACAGGA 486
DB 180 TTGGAAGGAAGAACTATGTCACAGCCAAACATTTGATATCTGACTCAGACAGGA 239
QY 487 AAGCACTTCATTTT 500
DB 240 AAGCACTTCATGTT 253

Search completed: October 10, 2002, 15:15:40
Job time : 90 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 : Search time 18.7333 Seconds

(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101a-9_COPY_1_500

Perfect score: 500
1 gaattctgactgcacagc.....cgatatgattcagtgatg 500

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents: NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	29.8	6.0	1541	US-08-200-232-3	Sequence 3, Appl
C 2	29.8	6.0	1541	PCT-US95-02219-3	Sequence 3, Appl
C 3	29.8	6.0	1541	PCT-US95-02219A-3	Sequence 1, Appl
C 4	29.8	6.0	3590	US-08-587-889-1	Sequence 1, Appl
C 5	29.8	6.0	3590	PCT-US96-09193-1	Sequence 1, Appl
C 6	29.6	5.9	4495	US-09-417-822-7	Sequence 6, Appl
C 7	29.6	5.9	4534	US-09-417-822-6	Sequence 7, Appl
C 8	29.6	5.9	4535	US-09-417-822-8	Sequence 8, Appl
C 9	29.6	5.9	5117	PCT-US95-05512-1	Sequence 1, Appl
C 10	29.6	5.9	5117	US-09-417-822-1	Sequence 1, Appl
C 11	29.6	5.9	663	US-08-463-115-89	Sequence 89, Appl
C 12	29.4	5.9	663	US-08-463-115-88	Sequence 88, Appl
C 13	29.4	5.9	669	US-08-463-115-88	Sequence 88, Appl
C 14	29.4	5.9	669	US-08-465-388-88	Sequence 88, Appl
C 15	29.4	5.9	669	US-08-375-709-8	Sequence 8, Appl
C 16	28.6	5.7	864	US-08-752-929-8	Sequence 8, Appl
C 17	28.6	5.7	37895	US-08-375-709-1	Sequence 1, Appl
C 18	28.6	5.7	37895	US-08-752-929-1	Sequence 1, Appl
C 19	28.6	5.7	37895	US-09-090-793-1	Sequence 1, Appl
C 20	28.6	5.7	1554	US-08-370-975B-10	Sequence 10, Appl
C 21	28.4	5.7	3494	US-09-139-802-200	Sequence 200, Appl
C 22	28.4	5.7	26764	US-08-370-975B-10	Sequence 10, Appl
C 23	28.4	5.7	3226	US-08-870-126-10	Sequence 10, Appl
C 24	27.8	5.6	2020	PCT-US95-13975-1	Sequence 1, Appl
C 25	27.8	5.6	6176	US-08-911-321-6	Sequence 6, Appl
C 26	27.8	5.6	6176	US-08-911-321-6	Sequence 6, Appl
C 27	27.6	5.5	2232	US-08-196-989B-3	Sequence 3, Appl

C 28	27.6	5.5	2232	US-08-760-936-3	Sequence 3, Appl
C 29	27.6	5.5	3516	US-09-188-930-257	Sequence 257, Appl
C 30	27.6	5.5	3706	US-08-207-904-16	Sequence 16, Appl
C 31	27.6	5.5	7305	US-08-286-740-4	Sequence 4, Appl
C 32	27.6	5.5	7305	PCT-US95-09576-4	Sequence 4, Appl
C 33	27.4	5.5	1673	US-08-294-143-6	Sequence 4, Appl
C 34	27.4	5.5	1673	US-08-256-331-4	Sequence 4, Appl
C 35	27.4	5.5	4560	US-08-256-703-1	Sequence 1, Appl
C 36	27.4	5.5	5635	US-08-136-742A-3	Sequence 3, Appl
C 37	27.4	5.5	5635	US-09-248-026-3	Sequence 3, Appl
C 38	27.4	5.5	5635	PCT-US93-11667-3	Sequence 3, Appl
C 39	27.4	5.5	6126	US-08-951-912-3	Sequence 3, Appl
C 40	27.4	5.5	6126	US-09-174-077-3	Sequence 3, Appl
C 41	27.4	5.5	6129	US-07-637-621-1	Sequence 1, Appl
C 42	27.4	5.5	6129	US-08-136-742A-1	Sequence 1, Appl
C 43	27.4	5.5	6129	US-08-135-809A-1	Sequence 1, Appl
C 44	27.4	5.5	6129	US-08-951-912-1	Sequence 1, Appl
C 45	27.4	5.5	6129	US-08-951-912-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-200-232-3/C
Sequence 3, Application US/08200232
Patent No. 5721349
GENERAL INFORMATION:
APPLICANT: Cover, Timothy L.
TITLE OF INVENTION: VACCINATING TOXIN-DEFICIENT H. PYLORI
TITLE OF INVENTION: AND RELATED METHODS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG P.C.
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/200,232
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 2200.023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1541
US-08-200-232-3

Query Match 6.0%; Score 29.8; DB 1; Length 1541;
Best Local Similarity 56.7%; Pred.No. 5.2;
Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 168 TGAATTGATGCCAGCGTTGGTGGCTGCTGACTTCTTGTGGTATGGGATGTCAT 227

[illegible]

RESULT 2

PCT-US95-02219-3/C
Sequence 3, Application PC/TUS9502219
GENERAL INFORMATION:
APPLICANT: Cover, Timothy L.
APPLICANT: Blaser, Martin J.
TITLE OF INVENTION: VACCINATING TOXIN-DEFICIENT H. PYLORI
TITLE OF INVENTION: AND RELATED METHODS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG P.C.
STREET: 127 Peachtree Street, Suite 1200
City: Atlanta
STATE: Georgia
COUNTRY: USA
Zip: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02219
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 2200.023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1541
PCT-US95-02219-3

Query Match	6.0%;	Score 29.8;	DB 5;	Length 1541;
Best Local Similarity	56.7%;	Pred. No. 5.2;		
Matches 55; Conservative	0;	Mismatches 42;	Indels 0;	Gaps 0;

OY	168	TGATGTATTACCAAGCGCTTGGTCGCCTACACTTTCTTGCTGTATGGGATGTAT	227
Dd	1399	TGTTGTGTTAAAGGCTAGCGCCTTTGTAATGCTCTTGCAATGTGTTGTCCTTG	1340
OY	228	CTCAAACTCATATTCACCACGAGCCCTTAGSAAAC	264
Dd	1339	CAGAATAATGTCATTAAGCGCTTGGACATCCAC	1303

RESULT 3

```

PCT-US95-02219A-3/c
: Sequence 3, Application PC/TUS9502219A
:
: GENERAL INFORMATION:
:
: APPLICANT: Cover, Timothy L.
:
: APPLICANT: Tammuray, Murali KR
:
: APPLICANT: Cao, Ping
:
: APPLICANT: Thompson, Stuart A.

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	6.0%;	Score 29.8;	DB 5;	Length 1541;
Best Local Similarity	56.7%;	Pred. No. 5.2;		
Matches 55; Conservative	0;	Mismatches 42;	Indels 0;	Gaps 0;

[illegible]

RESULT 4

US-08-587-889-1
Sequence 1, Application US/08587889
Patent No. 5654397
GENERAL INFORMATION:
APPLICANT: CAO, Zhaodan
APPLICANT: CROSTON, Glenn E.
APPLICANT: GOEDEL, David V.
TITLE OF INVENTION: INTERLUKIN-1 RECEPTOR-ASSOCIATED
TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSES: FLEHR, HOBACH, TEST, ALBRITTON & HERBERICH
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

; TITLE OF INVENTION: ATR-2
; FILE REFERENCE: 27866/35633

```

CURRENT APPLICATION NUMBER: US/09/417,822
CURRENT FILING DATE: 1999-10-14
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 7
LENGTH: 4534
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: pmtw-312r-5
US-09-417-822-7

Query Match 5.9%; Score 29.6; DB 4; Length 4534;
Best Local Similarity 45.6%; Pred. No. 10;
Matches 104; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

62 AAATCAAGACATTTTACGAAGTCAAGTGGCCGCCCTCATCTCTCATGTGCTTC 121
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 315 AAAGGATTTGATTTCAACAGTTTCCAAATGCTTGTGGAATGAGATTCACAGTGCACACT 256
122 GAATTAATTCATCAGAGCTCTATGATGATGAGATGCTTCCTCGATGTTGATGCCA 181
125 AGTTGAACACGGCAACATCGACACATCTCTGTAAAGATCATCTGGCAAGAGAGGAC 196
182 AGGCTTTGGTGGCTGCTGATCTTCTGTGTATGGGATGTCATCTCAACATCAATA 241
195 AGAGCATGTAGACGTCTGATGCTCCCTGACAGACATATTTCATCTTCAACAGAGTGTG 136
242 TCACGACAGCCCTTGAGAACAGGTTGAGACGAGCTAGAAAAA 289
135 GCTAAAGGCTTTGCAAAATTTATGAAAGACGAGAGAGAAATATA 88

RESULT 8

US-09-417-822-8/c
Sequence 8, Application US/09417822
Patent No. 6344549
GENERAL INFORMATION:
APPLICANT: Keegan, Kathy
TITLE OF INVENTION: ATR 2
FILE REFERENCE: 27866/3583
CURRENT APPLICATION NUMBER: US/09/417,822
CURRENT FILING DATE: 1999-10-14
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 8
LENGTH: 4535
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: pmtw-312r-7
US-09-417-822-8

Query Match 5.9%; Score 29.6; DB 4; Length 4535;
Best Local Similarity 45.6%; Pred. No. 10;
Matches 104; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

62 AAATCAAGACATTTTACGAAGTCAAGTGGCCGCCCTCATCTCTCATGTGCTTC 121
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 315 AAAGGATTTGATTTCAACAGTTTCCAAATGCTTGTGGAATGAGATTCACAGTGCACACT 256
122 GAATTAATTCATCAGAGCTCTATGATGATGAGATGCTTCCTCGATGTTGATGCCA 181
125 AGTTGAACACGGCAACATCGACACATCTCTGTAAAGATCATCTGGCAAGAGAGGAC 196
182 AGGCTTTGGTGGCTGCTGATCTTCTGTGTATGGGATGTCATCTCAACATCAATA 241
195 AGAGCATGTAGACGTCTGATGCTCCCTGACAGACATATTTCATCTTCAACAGAGTGTG 136
242 TCACGACAGCCCTTGAGAACAGGTTGAGACGAGCTAGAAAAA 289
135 GCTAAAGGCTTTGCAAAATTTATGAAAGACGAGAGAGAAATATA 88

RESULT 9
US-08-854-585-1/c
Sequence 1, Application US/08854585
Patent No. 6114140
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K. and Stman, Arne
TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, Suite 6300
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,585
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/237,940
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27866/31954
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5117 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 350..4364
US-08-854-585-1

Query Match 5.9%; Score 29.6; DB 3; Length 5117;
Best Local Similarity 57.6%; Pred. No. 11;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

85 TCAAGTGTGGCCGCCCTCATCTCTCATGATGTTGGAATTAATTCATCAGAGCTTAT 144
111111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 2295 TCAAGTGTGGCCGCCCTCATCTCTCATGATGTTGGAATTAATTCATCAGAGCTTAT 176
145 CGATCAGTGGGAGATGCTCTCCGATGTTGA 176
2235 ACATTCGTGGCCGCTGTCTGCTGCTGCA 2204

RESULT 10

PCT-US95-05512-1/c
Sequence 1, Application PC/US9505512
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K. and Stman, Arne
TITLE OF INVENTION: Density Enhanced Protein Tyrosine
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray &
STREET: 233 South Wacker Drive, Suite 6300
CITY: Chicago
STATE: Illinois

Db	567	ACGTA	CTGGGA	ATGAAG	ATCCAA	CTTCTT	CACAG	CACTGT	CACTGT	CGAGATG	508
Qy	218	GGGAT	GCATCT	CAAAACAT	236						
Db	507	AGGAA	CTTCC	ACAGACAT	489						

```

RESULT 13
US-08-465-388-89/C
: Sequence 89, Application US/08465388
: Patent No. 5753488
:
: GENERAL INFORMATION:
: APPLICANT: WILLIAM JOHN MARTIN
: TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
: TITLE OF INVENTION: AND RELATED VACCINES
: NUMBER OF SEQUENCES: 104
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: MEDIUM TYPE: Storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Fastseq Version 1.5
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/465,388
: FILING DATE: June 5, 1995
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: including application
: PRIOR APPLICATION DATA: described below:
:
: APPLICATION NUMBER: 08/157,811
: FILING DATE: NO.5753488number23, 1993
: APPLICATION NUMBER: 07/887,502
: FILING DATE: May 22, 1992
: APPLICATION NUMBER: 07/704,814
: FILING DATE: May 23, 1991
: APPLICATION NUMBER: 07/766,039
: FILING DATE: September 20, 1991
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 213/300
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
:
: TELEX: 67-3510
:
: INFORMATION FOR SEQ. ID NO.: 89:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 663 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-465-388-89

```

Best Local Match	5.9%;	Score 29.4;	DB 1;	Length 663;
Query Match Similarity	60.8%;	Pred. No. 4.6;		
Matches	48;	Conservative	0;	Mismatches 31; Indels 0; Gaps 0
QY	158	ATGTCCTCCGCGATGTCGACGAAGCCCTTGCGCCCTGCATCTTTCGCGCATG	217	
Db	567	ACGTCACCTGGGAGATGGAAGTCACATTTCTTCACACACACTGTGCTGCTGCGAGATG	508	
QY	218	GGGATGTCATCTCAACAT	236	
Db	507	AGGAACCTCTCCACAGACAT	489	

US-08-463-115-88/c
Sequence 88, Application US/08463115
Patent No. 5703221

GENERAL INFORMATION:
APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
TITLE OF INVENTION: AND RELATED VACCINES
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Fastseq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,115
FILING DATE: June 5, 1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/157,811
FILING DATE: No. 5703221,ember 23, 1993
APPLICATION NUMBER: 07/687,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 213/301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ. ID NO.: 88:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-463-115-88

```

Query Match Similarity      5.9%; Score 29.4; DB 1; Length 669;
Best local similarity      60.8%; Pred. No. 4.6;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0

QY 158 ATGTCCTCCGGATGGATGGATGCCAGGCTTTGGCCGCTGACTTCTTGGTGATG 217
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 ACCTACTCTGGGAGATGAAGATCCACATCTTCACAGCAGCTGACTGCTGGAGGATG 514

QY 218 GGGATGTCATCTCAACAT 236
      ||| || ||| |||
Db 513 AGGAACTCCACAGACAT 495

RESULT 15
US-08-465-388-88/c
; Sequence 88, Application US/08465388
; Patent No. 3753488
; GENERAL INFORMATION:

```


APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
TITLE OF INVENTION: AND RELATED VACCINES
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,388
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/157,811
FILING DATE: No. 5753488ember 23, 1993
APPLICATION NUMBER: 07/887,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Waidburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 213/300
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 88:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-465-388-88

Query Match 5.9% Score 29.4; DB 1; Length 669;
Best local Similarity 60.8%; Pred. No. 4.6;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
OY 158 ATGTCTCCGATGATGCGCAAGGCTTGGTGGCGCTGACCTTCTGTGTATG 217
DB 573 ACCTACTGGAGATGAAAGTCCACATCTCTACACGACCTGACTGCTGCGAGATG 514
OY 218 GGGATGTCATCTCAACAT 236
DB 513 AGGAAGCTTCACAGACAT 495

Search completed: October 10, 2002, 18:14:28
Job time : 38.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 695.733 Seconds

(without alignments)
9699.805 Million cell updates/sec

Title: us-09-489-101a-9_copy_1_500

Perfect score: 500
Sequence: 1 gaattcctgactgcacagc.....cgatgatcttactgacgacg 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : EST:*

1: em_estha:*

2: em_esthum:*

3: em_estlin:*

4: em_estlun:*

5: em_estlun:*

6: em_estlun:*

7: em_estlun:*

8: em_hic:*

9: gb_estl:*

10: gb_estl:*

11: gb_hic:*

12: gb_ges:*

13: em_ges_hum:*

14: em_ges_inv:*

15: em_ges_pln:*

16: em_ges_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	713	10	BE384912 601276436
2	500	100.0	872	9	AL549103 AL549103
3	500	100.0	928	9	AL556496 AL556496
4	500	100.0	1057	9	AL549090 AL549090
5	499.6	99.9	865	9	AL527428 AL527428
6	499.2	99.8	976	9	AL532402 AL532402
7	498.4	99.7	1138	10	BM464486 BM464486
8	493.2	98.6	920	9	AL518699 AL518699
9	489	97.6	837	10	BE478628 BE478628
10	488	97.6	743	10	BE408507 BE408507
11	487.4	97.5	700	9	AL553746 AL553746
12	486.4	97.3	690	10	BP976694 BP976694
13	486.4	97.3	887	10	BE685786 BE685786
14	485	97.0	936	10	BT753143 BT753143
15	484.8	97.0	955	10	BE740422 BE740422
16	483.8	96.8	1014	10	BE740422 BE740422
17					BE476370 BE476370

18	477	95.4	876	10	BE386238
19	476.4	95.3	699	10	BM008455
20	476	95.2	712	10	BF307478
21	476	95.2	798	10	BE272120
22	474.4	94.9	724	10	BE478857
23	473.8	94.8	866	10	BE409574
24	469.8	94.0	799	10	BE409574
25	462.6	92.5	870	10	BE476233
26	462	92.4	999	10	BM464540
27	460.4	92.1	752	10	BM083375
28	454	90.8	860	10	BE282087
29	453.2	90.6	797	10	BE902069
30	440.2	88.0	757	10	BE337550
31	438	87.6	859	10	BE325760
32	437.2	87.4	805	10	BE150893
33	436.8	87.4	679	10	BE620903
34	429	85.8	813	10	BE340922
35	423.8	84.8	698	10	BE409885
36	420	84.0	735	10	BE602244
37	414.6	82.9	862	10	BE475925
38	405.2	81.0	930	10	BE600224
39	397.2	79.4	722	10	BE874220
40	395.8	79.2	703	10	BE158553
41	395	79.0	892	10	BE312821
42	391	78.2	661	10	BE314366
43	386.4	77.3	728	10	BE310996
44	382.8	76.6	866	10	BE310996
45	379.4	75.9	1058	10	BE420072

ALIGNMENTS

RESULT 1
BE384912 713 bp mRNA linear EST 21-JUL-2000
LOCUS 601276436F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3617337 5',
DEFINITION mRNA sequence.
ACCESSION BE384912
VERSION BE384912
KEYWORDS BE384912.1 GI:9330277
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 713)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DMP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM284 row: h column: 10
High quality sequence stop: 611.

FEATURES

1..713
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3617337"
/clone_id="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site: 1; XhoI; Site: 2;
EcoRI: cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAC(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of

California, Berkeley) using 2AP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 186 a 162 c 170 g 195 t

ORIGIN

Query Match 100.0%; Score 500; DB 10; Length 713;
Best Local Similarity 100.0%; Pred. No. 2,38-144;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCCTGACGCGCAGAGGTGACAGAAACATTTGCTTTGTTGCTGGAAGCTGCT 60
|||||
32 GAATTCCTGACGCGCAGAGGTGACAGAAACATTTGCTTTGTTGCTGGAAGCTGCT 91
Db
61 CAATCAAGAACATTTACTGAAATCAAGTGGCCGCCCTACATCTCTCAATGCTGTT 120
|||||
92 CAATCAAGAACATTTACTGAAATCAAGTGGCCGCCCTACATCTCTCAATGCTGTT 151
Db
121 CGAATATATTCATCAGAGCTCTATCGATCAGTGGAGATGCTCCGCGATGCTGATGCC 180
|||||
152 CGAATATATTCATCAGAGCTCTATCGATCAGTGGAGATGCTCCGCGATGCTGATGCC 211
Db
181 AAGCTTTGCTGGCTCTGACCTCTCTCTGCTGATGGGATGCTCAATCAATCAAT 240
|||||
212 AAGCTTTGCTGGCTCTGACCTCTCTCTGCTGATGGGATGCTCAATCAATCAAT 271
Db
241 ATCACCAGAGCCCTTGAGGAAACAGAGTTGAGACGAAAGCTGAAAAAATGTTCTGTG 300
|||||
272 ATCACCAGAGCCCTTGAGGAAACAGAGTTGAGACGAAAGCTGAAAAAATGTTCTGTG 331
Db
301 ATGACGATGATCTTCAAGGATGATCCCGACGACCAACCTGCTGACAGAAACAT 360
|||||
332 ATGACGATGATCTTCAAGGATGATCCCGACGACCAACCTGCTGACAGAAACAT 391
Db
361 GTGGTAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
|||||
392 GTGGTAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 451
Db
421 CTCGGCGTTTGCATTTCTCTGAGCTGTTTCAAGGCAATGATGATGATGATGATGATGAT 480
|||||
452 CTCGGCGTTTGCATTTCTCTGAGCTGTTTCAAGGCAATGATGATGATGATGATGATGAT 511
Db
481 CGATATGATTTACTGATG 500
|||||
512 CGATATGATTTACTGATG 511
Db

RESULT 2
AL549103 872 bp mRNA linear EST 16-FEB-2001
LOCUS AL549103 LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1051YK06 5
DEFINITION prime, mRNA sequence.
ACCESSION AL549103
VERSION AL549103.1 GI:12884761
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 872)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source
1..872
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1051YK06"
/clone.lib="LTI.NFL006.PL2"
/tissue_type="placenta"

BASE COUNT 203 a 212 c 236 g 216 t 5 others

ORIGIN

Query Match 100.0%; Score 500; DB 9; Length 872;
Best Local Similarity 100.0%; Pred. No. 2,58-144;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCCTGACGCGCAGAGGTGACAGAAACATTTGCTTTGCTGGAAGCTGCT 60
|||||
246 GAATTCCTGACGCGCAGAGGTGACAGAAACATTTGCTTTGCTGGAAGCTGCT 305
Db
61 CAATCAAGAACATTTACTGAAATCAAGTGGCCGCCCTACATCTCTCAATGCTGTT 120
|||||
306 CAATCAAGAACATTTACTGAAATCAAGTGGCCGCCCTACATCTCTCAATGCTGTT 365
Db
121 CGAATATATTCATCAGAGCTCTATCGATCAGTGGAGATGCTCCGCGATGCTGATGCC 180
|||||
366 CGAATATATTCATCAGAGCTCTATCGATCAGTGGAGATGCTCCGCGATGCTGATGCC 425
Db
181 AAGCTTTGCTGGCTCTGACCTCTCTCTGCTGATGGGATGCTCAATCAATCAAT 240
|||||
426 AAGCTTTGCTGGCTCTGACCTCTCTCTGCTGATGGGATGCTCAATCAATCAAT 485
Db
241 ATCACCAGAGCCCTTGAGGAAACAGAGTTGAGACGAAAGCTGAAAAAATGTTCTGTG 300
|||||
486 ATCACCAGAGCCCTTGAGGAAACAGAGTTGAGACGAAAGCTGAAAAAATGTTCTGTG 545
Db
301 ATGACGATGATCTTCAAGGATGATCCCGACGACCAACCTGCTGACAGAAACAT 360
|||||
546 ATGACGATGATCTTCAAGGATGATCCCGACGACCAACCTGCTGACAGAAACAT 605
Db
361 GTGGTAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
|||||
606 GTGGTAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 665
Db
421 CTCGGCGTTTGCATTTCTCTGAGCTGTTTCAAGGCAATGATGATGATGATGATGATGAT 480
|||||
666 CTCGGCGTTTGCATTTCTCTGAGCTGTTTCAAGGCAATGATGATGATGATGATGATGAT 725
Db
481 CGATATGATTTACTGATG 500
|||||
726 CGATATGATTTACTGATG 745
Db

RESULT 3
AL556496 928 bp mRNA linear EST 16-FEB-2001
LOCUS AL556496 LTI.NFL006.PL2 Homo sapiens cDNA clone CS0K0061A17 5
DEFINITION prime, mRNA sequence.
ACCESSION AL556496
VERSION AL556496.1 GI:12899223
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 928)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-0190(5') primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@life.com url : http://fulllength.invitrogen.com"

FEATURES

Email: segregre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Location/Qualifiers

source

```
1. .928
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDK006YK17"
/clone_1fb="url.NF1006.Pl2"
/tissue_type="placenta"
/notes="vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng liang life technologies,
a division of invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
```

218 a 224 c 233 g 229 t 4 others

Query Match	100.0%;	Score 500;	DB 9;	Length 928;
Best Local Similarity	100.0%;	Pred. No. 2.6e-144;		
Matches 500;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 GAATTCCTGACGCGCCACAGGCTTACAGSAAACATTTGCTTTTGTCTGGAAAGCTGCT 60
|||||
Db 256 GAATTCCTGACGCGCCACAGGCTTACAGSAAACATTTGCTTTTGTCTGGAAAGCTGCT 315

Qy 61 CAATCAAGAAGACATTTACTGAAGTCAAAGTGGTGGCCGCCCTACATCTCATGTGGTT 120
|||||
|||
Db 316 CAATCAAGAAGACATTTACTGAAGTCAAAGTGGTGGCCGCCCTACATCTCTCAATGTGGTT 375

0y 121 CGAATATTACATCAGAGCTCATCGATCACATGGGAGTGTCTCCCGATGTTGATGCC 180
|||||
Db 376 CGAATATTACATCAGAGCTCATCGATCACATGGGAGATGTCCTCCCGATGTTGATGCC 435

436 AAGCTTTGGTCGCTGACTTCTCTGTGATGGGATGTCATCTCAACATCAAT 495

ATCACCAGACCCCTTGAGGACACACAGCTTGAGACGGAACTAGAAAAAATGTTCCTGTG 555

5 |||
Db 556 ATGACGATGATCTTCAAGGAGTCATCCCGCAGCACCACCAACTGTTGGCAGAGACAAAT 615
361 GCGCTACTGCGCTGTCGATAGTATCCACAAACAAGCGCTTCCTCAATTTTCACAGAGACCCAGCGT 420

421	CTCCGGCGTTTTCATTTCCCTTGAGGCGTGTTCAGGGGAGTGTGATGATGAGTGGAGGTT	480
Db	616 GTGGTAGTGGCTGTGGATAGTACCAACAACAGGGTTCCTCATTTTCAAGAAGACCAGGGT	675
Oy	421	

Db 676 CTCCGGCGTTTGCATTTCTCTGAGCCTGTTTCAGGCGAGTAGTGAGAGGTT 735
 Oy 481 CGATATGATTACTGGAATG 500

Db 736 CGATATGATTACTGATG 755

AL549090	1057 bp	mRNA	linear	EST 16-FEB-2001
LOCUS				
DEFINITION	AL549090 LTI_NFL006_PL2 Homo sapiens CDNA, clone CSODI051YB12.5			

ACCESSION	AL549090
VERSION	AL549090.1
KEYWORDS	GI:12884735
SOURCE	EST.
ORGANISM	human

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 1057)
Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre national de séquençage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Location/Qualifiers

```
1.1057
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI051YB12"
```

```

/tissue_type="Placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st stru
was primed with a NotI-oligo(dT) primer. Five pu
rified double-stranded cDNA was digested with

```

cloned into the Not I and Eco RV sites of the pCMV vector. Library was normalized. Library was constructed. Contact: Feng Liang Life Tech, a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610
Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com>"

ORIGIN	Score	DB	Length
Query Match	100.0%	500	9
Best Local Similarity	100.0%		
	Pred. No.	2.7e-144	

Matches	500;	Conservative	0;	Mismatches	0;	Indels	0;
Oy	1	GAATTCCTGACTGCACAGGTGTACAGAAACATTTGCTTTTGTTCCTGGAAGCTGCT					

DU 217 GAAATTCCTGACAGCCACAGGGTGTATCAGGGAAACAATTTGTAATTGTAATCGGCAATGC

OY 61 CAATATCAAGAACAATTACTGTGAAGTCAAAGTGTCGCCGCCCTACATCTCTCAAATGTGGTTTT
|||||
|||

287

121 CGAATATTACATCGAGCTCTATCGATCAGTGGGAGATGTCCTCGGTGATGTGGATGCC
 337 CGAATATTACATCGAGCTCTATCGATCAGTGGGAGATGTCCTCGGTGATGTGGATGCC

Oy 181 AAGCATTGTCGCGCTGCACCTTCTCTGTCGTATGGGATGTCATCTCAACATCAAT
 |||||||
 Db 397 AAGCTTTGTGGGCTTCACCTTCTCTGTCGTATGGGATGTCATCTCAACATCAAT
 |||||||

QY 241 ATCACCAGAGCCCTTGAGGACACACAGTTGAGACGGAGCTAGAAAAAATGTTTCTGTG
 457 ATCACCAGAGCCCTTGAGGACACACAGTTGAGACGGAGCTAGAAAAAATGTTTCTGTG
 DB

QY 301 ATGACGATGATCTTCAAGSAGTCATCCCCCAGCCACCCTGTTGCCAGAGACAAAT
|||||
Db 517 ATGACGATGATCTTCAAGSAGTCATCCCCCAGCCACCCTGTTGCCAGAGACAAAT

Oy	361 GTGGTAGTGGCTGTGATAGTACCACAACAGGGTTCTCATTTTCAGAAGACCAGGG
Dd	577 GTGGTAGTGGCTGTGATAGTACCACAACAGGGTTCTCATTTTCAGAAGACCAGGG

QY	481	CGATATGATTACTGATTC	500
Db	697	CGATATGATTACTGATTC	716

RESULT 5	865 bp	mrna	linear	EST 13
AL527428				
LOCUS				
AL527428				

LOCUS	AL527428	865 bp	mRNA	linear	EST 13-FEB-2001
-------	----------	--------	------	--------	-----------------

```

DEFINITION   AL527428 LfL_NFL003_NBC3 Homo sapiens cDNA clone CSDDC021YU03 5
PRIME, mRNA sequence.
ACCESSION    AL527428
VERSION      AL527428.1 GI:12790921
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 865)
AUTHORS      Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDC021YU03"
/clone_1ib="LfL_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com" 5 others

BASE COUNT   206 a 209 c 229 g 216 t
ORIGIN
Query Match          99.8%; Score 499.6; DB 9; Length 865;
Best Local Similarity 99.8%; Pred. No. 3.3e-144;
Matches 499; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGACGCGCAGCTGACAGAAACATTTGCTTTGTCGGAACCTGCT 60
    |||||||
Db 219 GAATTCCTGACGCGCAGCTGACAGAAACATTTGCTTTGTCGGAACCTGCT 278

QY 61 CAATCAAGACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCTCAATGTGTT 120
    |||||||
Db 279 CAATCAAGACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCTCAATGTGTT 338

QY 121 CGAATTAATTAATCAAGAGCTCTATGCATCACTGAGAGATGTCTCCCGGATGATGATGCC 180
    |||||||
Db 339 CGAATTAATTAATCAAGAGCTCTATGCATCACTGAGAGATGTCTCCCGGATGATGATGCC 398

QY 181 AAGGCTTTGGGCTGCTGACTTTCTTTGGTGTATGGGATGTCATCTCAACATCAAT 240
    |||||||
Db 399 AAGGCTTTGGGCTGCTGACTTTCTTTGGTGTATGGGATGTCATCTCAACATCAAT 458

QY 241 ATCACCAGAGCCCTTGAGAAACACAGTGTGAGACGGAACCTAGAAAAAATGTTCTGTG 300
    |||||||
Db 459 ATCACCAGAGCCCTTGAGAAACACAGTGTGAGACGGAACCTAGAAAAAATGTTCTGTG 518

QY 301 ATGACGATGATCTTCAAGAGGATATCCCGACGACCAACATCGTTGCCAGCAACATCAAT 360
    |||||||
Db 519 ATGACGATGATCTTCAAGAGGATATCCCGACGACCAACATCGTTGCCAGCAACATCAAT 578

QY 361 GTGGTAGTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
    |||||||
Db 579 GTGGTAGTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638

QY 421 CTCGGGCTTTGCTGATTTCTGTGAGCTGTGTTCAAGGAGTAGAGATGAGATGAGGTT 480
    |||||||

```

```

Db 639 CTCGGGCTTTGCTGATTTCTGTGAGCTGTGTTCAAGGAGTAGAGATGAGATGAGGAGTT 698
QY 481 CGATATGATTTACTGTGATG 500
    |||||||
Db 699 CGATATGATTTACTGTGATG 718

RESULT 6
AL532402
LOCUS       AL532402 976 bp mRNA linear EST 13-FEB-2001
DEFINITION AL532402 LfL_NFL001_NBC4 Homo sapiens cDNA clone CSDDM014YD01 5
PRIME, mRNA sequence.
ACCESSION  AL532402
VERSION     AL532402.1 GI:12795895
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 976)
AUTHORS      Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDM014YD01"
/clone_1ib="LfL_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com" 10 others

BASE COUNT   228 a 237 c 261 g 240 t
ORIGIN
Query Match          99.8%; Score 499.2; DB 9; Length 976;
Best Local Similarity 99.6%; Pred. No. 4.6e-144;
Matches 498; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTGACGCGCAGCTGACAGAAACATTTGCTTTGTCGGAACCTGCT 60
    |||||||
Db 253 GAATTCCTGACGCGCAGCTGACAGAAACATTTGCTTTGTCGGAACCTGCT 312

QY 61 CAATCAAGACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCTCAATGTGTT 120
    |||||||
Db 313 CAATCAAGACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCTCAATGTGTT 372

QY 121 CGAATTAATTAATCAAGAGCTCTATGCATCACTGAGAGATGTCTCCCGGATGATGATGCC 180
    |||||||
Db 373 CGAATTAATTAATCAAGAGCTCTATGCATCACTGAGAGATGTCTCCCGGATGATGATGCC 432

QY 181 AAGGCTTTGGGCTGCTGACTTTCTTTGGTGTATGGGATGTCATCTCAACATCAAT 240
    |||||||
Db 433 AAGGCTTTGGGCTGCTGACTTTCTTTGGTGTATGGGATGTCATCTCAACATCAAT 492

QY 241 ATCACCAGAGCCCTTGAGAGACACAGTGTGAGACGGAACCTAGAAAAAATGTTCTGTG 300
    |||||||
Db 493 ATCACCAGAGCCCTTGAGAGACACAGTGTGAGACGGAACCTAGAAAAAATGTTCTGTG 552

```

OY 301 ATGACGATGATCTTCAAGAGATGATCCCGACCCAGCACTGCTGTCACGACAGCAT 360
 DB 553 ATGACGATGATCTTCAAGAGATGATCCCGACCCAGCACTGCTGTCACGACAGCAT 612
 OY 361 GTGGTAGTGCTGTGTAGTACCAACAAAGAGGTTCTCCATTCTTCAAGAGAGCCAGGCT 420
 DB 613 GTGGTAGTGCTGTGTAGTACCAACAAAGAGGTTCTCCATTCTTCAAGAGAGCCAGGCT 672
 OY 421 CTCGGCGGTTTTGCATTTCTCTGAGCTGTTTTCAGGCGAGTGTAGTGTAGTGTGAGT 480
 DB 673 CTCGGCGGTTTTGCATTTCTCTGAGCTGTTTTCAGGCGAGTGTAGTGTAGTGTGAGT 732
 OY 481 CGATATGATTTACTGATG 500
 DB 733 CGATATGATTTACTGATG 752

RESULT 7
 BM464486 1138 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT_6439009 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535574
 DEFINITION 5', mRNA sequence.
 ACCESSION BM464486
 VERSION BM464486.1 GI:18513528
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1138)
 NIH_MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://lmage.llnl.gov
 Plate: LLM12223 row: n column: 23
 High quality sequence start: 17
 High quality sequence stop: 650.

FEATURES
 Location/Qualifiers
 1..1138
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5535574"
 /clone_id="NIH_MGC_71"
 /tissue_type="leiomysarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pCMV-SPORT6; site_1: NotI;
 site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2.1 kb."
 BASE COUNT 256 a 310 c 298 g 274 t

ORIGIN
 Query Match 99.7%; Score 498.4; DB 10; Length 1138;
 Best local Similarity 99.8%; Pred. No. 8,6e-144;
 Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 DB 288 GAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347
 OY 61 CAATTCACAAACATTTACTGAAAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 DB 348 CAATTCACAAACATTTACTGAAAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
 OY 121 CGAATATTCATGAGCTGTATGATGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

DB 408 CGAATATTCATGAGCTGTATGATGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467
 OY 181 AAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 468 AAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
 OY 241 ATCCACAGGCCCTTGAGGACACAGGTTGAGACGGAAGCTGAGAAAAATGTTCTGTG 300
 DB 528 ATCCACAGGCCCTTGAGGACACAGGTTGAGACGGAAGCTGAGAAAAATGTTCTGTG 587
 OY 301 ATGACGATGATTTCAAGAGTATCCCGACGACGACCAACTGCTGTCACGACAGCAT 360
 DB 588 ATGACGATGATTTCAAGAGTATCCCGACGACGACCAACTGCTGTCACGACAGCAT 647
 OY 361 GTGGTAGTGCTGTGTAGTACCAACAAAGAGGTTCTCCATTCTTCAAGAGAGCCAGGCT 420
 DB 648 GTGGTAGTGCTGTGTAGTACCAACAAAGAGGTTCTCCATTCTTCAAGAGAGCCAGGCT 707
 OY 421 CTCGGCGGTTTTGCATTTCTCTGAGCTGTTTTCAGGCGAGTGTAGTGTAGTGTGAGT 480
 DB 708 CTCGGCGGTTTTGCATTTCTCTGAGCTGTTTTCAGGCGAGTGTAGTGTAGTGTGAGT 767
 OY 481 CGATATGATTTACTGATG 500
 DB 768 CGATATGATTTACTGATG 787

RESULT 8
 AL518699 920 bp mRNA linear EST 13-FEB-2001
 LOCUS AL518699 LTI_NF1011.NBC1 Homo sapiens cDNA clone CS0DA011YH01 5
 DEFINITION AL518699 LTI_NF1011.NBC1 Homo sapiens cDNA clone CS0DA011YH01 5
 PRIMER, mRNA sequence.
 ACCESSION AL518699
 VERSION AL518699.1 GI:12782192
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 920)
 Li, W.B., Gruber, C., Jessup, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL Contact: Genoscope
 COMMENT Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 Location/Qualifiers
 1..920
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DA011YH01"
 /clone_id="LTI_NF1011.NBC1"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library is not normalized, but is the control for
 the normalized libraries. Library was constructed by Life
 Technologies. Contact : Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifestech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 213 a 218 c 247 g 231 t 11 others

ORIGIN
 Query Match 98.6%; Score 493.2; DB 9; Length 920;
 Best local Similarity 98.4%; Pred. No. 3.3e-142;
 Matches 492; Conservative 5; Mismatches 3; Indels 0; Gaps 0;


```

OY 1 GAATTCCTGACCTCCACAGCTGTACAGGAACATTTGCTTTGTTGCTGGAAGCTGCT 60
DB 262 GAATTCCTGACCTCCACAGCTGTACAGGAACATTTGCTTTGTTGCTGGAAGCTGCT 321
OY 61 CAAATCAAGAACATTTACTGAGTCAAGAGTGGCGCCCTACATCTGCAATGCTGTT 120
DB 322 CAAATCAAGAACATTTACTGAGTCAAGAGTGGCGCCCTACATCTGCAATGCTGTT 381
OY 121 CGAATTAATTAATCAAGAGCTGTATGATCACTGGAGATGCTCTCGTGAATGTTGATGCC 180
DB 382 CGAATTAATTAATCAAGAGCTGTATGATCACTGGAGATGCTCTCGTGAATGTTGATGCC 441
OY 181 AAGCTTTGCTGCGCTCTGATCTTCTGCTGTATGATGAGGATGCTCAATCAATCAAT 240
DB 442 AAGCTTTGCTGCGCTCTGATCTTCTGCTGTATGATGAGGATGCTCAATCAATCAAT 501
OY 241 ATCACCAGAGCCCTTGAGGAGACAGAGTGTGAGAGCGAAGCAAAAAATGTTCTGTG 300
DB 502 ATCACCAGAGCCCTTGAGGAGACAGAGTGTGAGAGCGAAGCAAAAAATGTTCTGTG 561
OY 301 ATGACAGATGATCTGCAAGAGATATCCCGCAGCCAGCACTGCTGCGCAGAGCAAT 360
DB 562 ATGACAGATGATCTGCAAGAGATATCCCGCAGCCAGCACTGCTGCGCAGAGCAAT 621
OY 361 GGTGTAGTGTGCTGTGATGATGACCAACAGAGGTTCTCCATTTTCAGAAAGCCAGGTT 420
DB 622 GGTGTAGTGTGCTGTGATGATGACCAACAGAGGTTCTCCATTTTCAGAAAGCCAGGTT 681
OY 421 CTCGCGGCTTTGCTGATTTCTGCTGAGCTGTGTCAGAGGAGTGAAGTGAAGTGAAGT 480
DB 682 CTCGCGGCTTTGCTGATTTCTGCTGAGCTGTGTCAGAGGAGTGAAGTGAAGTGAAGT 741
OY 481 CGATATGATTTACTGATG 500
DB 742 CGATATGATTTACTGATG 761

RESULT 9
LOCUS Bg478628 837 bp mRNA linear EST 21-MAR-2001
DEFINITION 602525515F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4643907 5',
ACCESSION Bg478628
VERSION Bg478628.1 GI:13410907
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsb@remail.nih.gov
Tissue Procurement: ATCC/DCT/DMP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1414 row: n column: 04
High quality sequence stop: 792.
Location/Qualifiers
1. 837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4643907"
/clone_1db="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10b (phage-resistant)"

```

```

BASE COUNT 216 a 197 c 201 g 223 t
ORIGIN
Query Match 97.8%; Score 489; DB 10; Length 837;
Best Local Similarity 99.8%; Pred. No. 6,5e-141;
Matches 500; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
/Note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using
ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

OY 1 GAATTCCTGACCTCCACAGCTGTACAGGAACATTTGCTTTGTTGCTGGAAGCTGCT 60
DB 52 GAATTCCTGACCTCCACAGCTGTACAGGAACATTTGCTTTGTTGCTGGAAGCTGCT 111
OY 61 CAAATCAAGAACATTTACTGAGTCAAGAGTGGCGCCCTACATCTGCAATGCTGTT 120
DB 112 CAAATCAAGAACATTTACTGAGTCAAGAGTGGCGCCCTACATCTGCAATGCTGTT 171
OY 121 CGAATTAATTAATCAAGAGCTGTATGATCACTGGAGATGCTCTCGTGAATGTTGATGCC 180
DB 172 CGAATTAATTAATCAAGAGCTGTATGATCACTGGAGATGCTCTCGTGAATGTTGATGCC 231
OY 181 AAGCTTTGCTGCGCTCTGATCTTCTGCTGTATGATGAGGATGCTCAATCAATCAAT 240
DB 232 AAGCTTTGCTGCGCTCTGATCTTCTGCTGTATGATGAGGATGCTCAATCAATCAAT 291
OY 241 ATCACCAGAGCCCTTGAGGAGACAGAGTGTGAGAGCGAAGCAAAAAATGTTCTGTG 300
DB 292 ATCACCAGAGCCCTTGAGGAGACAGAGTGTGAGAGCGAAGCAAAAAATGTTCTGTG 351
OY 301 ATGACAGATGATCTGCAAGAGATATCCCGCAGCCAGCACTGCTGCGCAGAGCAAT 360
DB 352 ATGACAGATGATCTGCAAGAGATATCCCGCAGCCAGCACTGCTGCGCAGAGCAAT 411
OY 361 GGTGTAGTGTGCTGTGATGATGACCAACAGAGGTTCTCCATTTTCAGAAAGCCAGGTT 420
DB 412 GGTGTAGTGTGCTGTGATGATGACCAACAGAGGTTCTCCATTTTCAGAAAGCCAGGTT 471
OY 421 CTCGCGGCTTTGCTGATTTCTGCTGAGCTGTGTCAGAGGAGTGAAGTGAAGTGAAGT 480
DB 472 CTCGCGGCTTTGCTGATTTCTGCTGAGCTGTGTCAGAGGAGTGAAGTGAAGTGAAGT 531
OY 481 CGAT-ATGATTTACTGATG 500
DB 532 CGATATGATTTACTGATG 552

RESULT 10
LOCUS BE408507 743 bp mRNA linear EST 21-JUL-2000
DEFINITION 601303024F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3637852 5',
ACCESSION BE408507
VERSION BE408507.1 GI:9344957
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 743)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsb@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MCC clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLNLT at: image.llnl.gov
 Plate: L1C337 row: 0 column: 05
 High quality sequence stop: 625.
 Location/Qualifiers

FEATURES

source

1. 743
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3637852"
 /clone_lib="NIH_MCC_21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOT87; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Size-selected for
 average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using 2AP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 192 a 170 c 174 g 207 t
 ORIGIN

Query Match 97.6%; Score 488; DB 10; Length 743;
 Best Local Similarity 99.8%; Pred. No. 1.3e-140;
 Matches 499; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GAATTCCTGAGTCCACAGGTGTACAGAAACATTTGCTTTGCTGCGAAGCTGCT 60
 DB 32 GAATTCCTGAGTCCACAGGTGTACAGAAACATTTGCTTTGCTGCGAAGCTGCT 91
 OY 61 CAATCAAAACAACTTACTGAAATCAAGTGTGCGCCCTACATCTCTCAATGTGTT 120
 DB 92 CAATCAAAACAACTTACTGAAATCAAGTGTGCGCCCTACATCTCTCAATGTGTT 151
 OY 121 CGAATAATTCATCAGAGCTCTATCGATCAGTGGAGATGTCTCTCGATGTGATGCC 180
 DB 152 CGAATAATTCATCAGAGCTCTATCGATCAGTGGAGATGTCTCTCGATGTGATGCC 211
 OY 181 AAGGCTTGTGCGCTGTGACTTCTTCTGCTGATGTGGAGATGTATCTCAATCAAT 240
 DB 212 AAGGCTTGTGCGCTGTGACTTCTTCTGCTGATGTGGAGATGTATCTCAATCAAT 271
 OY 241 ATCCACGAGCCCTTGAGAACACAGGTTGAGACGGAACCTAGAAAAATGTTCTGTG 300
 DB 272 ATCCACGAGCCCTTGAGAACACAGGTTGAGACGGAACCTAGAAAAATGTTCTGTG 331
 OY 301 ATGACGATGATCTTCAAGAGATCATCCCGACGACCAACCTCGTGGCAGAGACAT 360
 DB 332 ATGACGATGATCTTCAAGAGATCATCCCGACGACCAACCTCGTGGCAGAGACAT 391
 OY 361 GTGGTAGTGGCTGTGATAGTACCAAGAGGTTTCCATTTTCAAGAACCCAGGT 420
 DB 392 GTGGTAGTGGCTGTGATAGTACCAAGAGGTTTCTCATTTTCAAGAACCCAG- 450
 OY 421 CTCGGCGCTTTTCATTTCTCTGAGCTGTCTTCAAGGACAGTAGTAGAGGTG 480
 DB 451 CTCGGCGCTTTTCATTTCTCTGAGCTGTCTTCAAGGACAGTAGTAGAGGTG 510
 OY 481 CGATATGATTTACTGATG 500
 DB 511 CGATATGATTTACTGATG 530

RESULT 11
 LOCUS AL553746 970 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL553746 L1_NFL006_PL2 Homo sapiens cDNA clone CS01080YF09 5
 prime, mRNA sequence.
 ACCESSION AL553746
 VERSION AL553746.1 GI:12893867
 KEYWORDS EST.

SOURCE

ORGANISM

human.
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 970)
 AUTHORS Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT

Contact: Genoscope
 Genoscope Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers

FEATURES

source

1. 970
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="CS01080YF09"
 /clone_lib="L1_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 was enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@life.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 222 a 246 c 257 g 240 t 5 others
 ORIGIN

Query Match 97.6%; Score 488; DB 9; Length 970;
 Best Local Similarity 99.8%; Pred. No. 1.4e-140;
 Matches 499; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GAATTCCTGAGTCCACAGGTGTACAGAAACATTTGCTTTGCTGCGAAGCTGCT 60
 DB 225 GAATTCCTGAGTCCACAGGTGTACAGAAACATTTGCTTTGCTGCGAAGCTGCT 284
 OY 61 CAATCAAAACAACTTACTGAAATCAAGTGTGCGCCCTACATCTCTCAATGTGTT 120
 DB 285 CAATCAAAACAACTTACTGAAATCAAGTGTGCGCCCTACATCTCTCAATGTGTT 344
 OY 121 CGAATAATTCATCAGAGCTCTATCGATCAGTGGAGATGTCTCTCGATGTGATGCC 180
 DB 345 CGAATAATTCATCAGAGCTCTATCGATCAGTGGAGATGTCTCTCGATGTGATGCC 404
 OY 181 AAGGCTTGTGCGCTGTGACTTCTTCTGCTGATGTGGAGATGTATCTCAATCAAT 240
 DB 405 AAGGCTTGTGCGCTGTGACTTCTTCTGCTGATGTGGAGATGTATCTCAATCAAT 464
 OY 241 ATCCACGAGCCCTTGAGAACACAGGTTGAGACGGAACCTAGAAAAATGTTCTGTG 300
 DB 465 ATCCACGAGCCCTTGAGAACACAGGTTGAGACGGAACCTAGAAAAATGTTCTGTG 523
 OY 301 ATGACGATGATCTTCAAGAGATCATCCCGACGACCAACCTCGTGGCAGAGACAT 360
 DB 524 ATGACGATGATCTTCAAGAGATCATCCCGACGACCAACCTCGTGGCAGAGACAT 583
 OY 361 GTGGTAGTGGCTGTGATAGTACCAAGAGGTTTCCATTTTCAAGAACCCAGGT 420
 DB 584 GTGGTAGTGGCTGTGATAGTACCAAGAGGTTTCTCATTTTCAAGAACCCAGGT 643
 OY 421 CTCGGCGCTTTTCATTTCTCTGAGCTGTCTTCAAGGACAGTAGTAGAGGTG 480
 DB 644 CTCGGCGCTTTTCATTTCTCTGAGCTGTCTTCAAGGACAGTAGTAGAGGTG 703
 OY 481 CGATATGATTTACTGATG 500
 DB 704 CGATATGATTTACTGATG 723

```

RESULT 12
LOCUS      Bf976694
DEFINITION 602243611f1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:434594 5',
            mRNA sequence.
ACCESSION  Bf976694
VERSION     Bf976694.1 GI:12343909
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   NIH-MGC http://mhc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS     Contact: Robert Strausberg, Ph.D.
            Email: cga@b-remail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1204 row: f column: 03
            High quality sequence stop: 665.
FEATURES
    source
        1..700
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:434594"
        /clone_1db="NIH_MGC_48"
        /tissue_type="Primary B-cells from tonsils (cell line)"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: B-cells; Vector: pORF7; Site_1: XhoI;
        Site_2: EcoRI; cDNA made by oligo-dT priming.
        Directionally cloned into EcoRI/XhoI sites using the
        following 5' adaptor: GGCACGAG(G). Size-selected >500bp
        for average insert size 1.8kb. Library constructed by Ling
        Hong in the laboratory of Gerald M. Rubin (University of
        California, Berkeley) using ZAP-cDNA synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies).
        Note: this is a NIH_MGC library."
BASE COUNT  170 a 173 c 172 g 185 t
ORIGIN
Query Match 97.5%; Score 487.4; DB 10; Length 700;
Best Local Similarity 99.6%; Pred. No. 1.9e-140;
Matches 499; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GAATTCCTGACTCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 60
    |||
DB 138 GAATTCCTGACTCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 197
QY 61 CAATATCAAAACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCAATGTGGTT 120
    |||
DB 198 CAATATCAAAACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCAATGTGGTT 257
QY 121 CGAATAATTCATCAGAGCTCTATCGATCTGCGAGATGTCCTCCGATGTGATGCC 180
    |||
DB 258 CGAATAATTCATCAGAGCTCTATCGATCTGCGAGATGTCCTCCGATGTGATGCC 317
QY 181 AAGGCTTTGGTGGCTCTGCTACTTTCTTCTGCTATGGGATGTCATCTCAACATCAAT 240
    |||
DB 318 AAGGCTTTGGTGGCTCTGCTACTTTCTTCTGCTATGGGATGTCATCTCAACATCAAT 377
QY 241 ATGACACAGAGCCCTTGAGGAACAGGTGAGAGAGAGCTGAAAAAAGTTTCTGTG 300
    |||
DB 378 ATGACACAGAGCCCTTGAGGAACAGGTGAGAGAGAGCTGAAAAAAGTTTCTGTG 437
QY 301 ATGACAGATGATCTTCAAGAGTATCCCCAGCCACCCCAACTGTTGCGACGAGACAT 360
    |||

RESULT 13
LOCUS      B685786
DEFINITION 602637891f1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:475463 5',
            mRNA sequence.
ACCESSION  B685786
VERSION     B685786.1 GI:13917183
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   NIH-MGC http://mhc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS     Contact: Robert Strausberg, Ph.D.
            Email: cga@b-remail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1624 row: f column: 24
            High quality sequence stop: 686.
FEATURES
    source
        1..690
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:475463"
        /clone_1db="NIH_MGC_48"
        /tissue_type="Primary B-cells from tonsils (cell line)"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: B-cells; Vector: pORF7; Site_1: XhoI;
        Site_2: EcoRI; cDNA made by oligo-dT priming.
        Directionally cloned into EcoRI/XhoI sites using the
        following 5' adaptor: GGCACGAG(G). Size-selected >500bp
        for average insert size 1.8kb. Library constructed by Ling
        Hong in the laboratory of Gerald M. Rubin (University of
        California, Berkeley) using ZAP-cDNA synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies).
        Note: this is a NIH_MGC library."
BASE COUNT  166 a 172 c 167 g 185 t
ORIGIN
Query Match 97.3%; Score 486.4; DB 10; Length 690;
Best Local Similarity 99.6%; Pred. No. 3.9e-140;
Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GAATTCCTGACTCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 60
    |||
DB 137 GAATTCCTGACTCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 196
QY 61 CAATATCAAAACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCAATGTGGTT 120
    |||
DB 197 CAATATCAAAACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCAATGTGGTT 256
    |||

```

QY 121 CGAATTAATTAACATGAGAGCTCTATCATGATCAGGAGATGCTCCCTGATGATGATGCC 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 257 CGAATTAATTAACATGAGAGCTCTATCATGATCAGGAGATGCTCCCTGATGATGATGCC 316
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 AAGGCTTTGGGCTGCTGATCTTCTGCTGATGAGGATGATGATGATGATGATGATGATGAT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 317 AAGGCTTTGGGCTGCTGATCTTCTGCTGATGAGGATGATGATGATGATGATGATGATGAT 376
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 ATACACAGAGCCCTTGAAGACACAGGTTGAGACGAGCTGAGAAAAAATGTTCTGTG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 377 ATACACAGAGCCCTTGAAGACACAGGTTGAGACGAGCTGAGAAAAAATGTTCTGTG 436
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 ATGAGATGATCTTCAAGAGATCATCCCGACGACCACTGCTGCTGACGAGAACAT 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 437 ATGAGATGATCTTCAAGAGATCATCCCGACGACCACTGCTGCTGACGAGAACAT 496
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 GGGGATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 497 GGGGATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 CTCGGGCTTTTGCATTTCTCTGAGCCTTTTCAGGACAGTAGATGATGATGATGATGATGAT 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 556 CTCGGGCTTTTGCATTTCTCTGAGCCTTTTCAGGACAGTAGATGATGATGATGATGATGAT 615
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 481 CGATATGATTTACTGATGATG 500
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 616 CGATATGATTTACTGATGATG 635

RESULT 14

B1753143 887 bp mRNA linear EST 25-SEP-2001
 LOCUS B1753143
 DEFINITION 60302599F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196531 5',
 mRNA sequence.

ACCESSION B1753143
 VERSION B1753143.1 GI:15744721
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M1491 row: p column: 24
 High quality sequence stop: 809.
 Location/Qualifiers

FEATURES

1. 887
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5196531"
 /clone_id="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

BASE COUNT 220 a 227 c 209 g 231 t
 ORIGIN
 Query Match 97.3%; Score 486.4; DB 10; Length 887;
 Best Local Similarity 99.6%; Pred. No. 4.3e-140;
 Matches 498; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GAATTCGACAGCTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 117 GAATTCGACAGCTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 176
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CAATCAAGAAACATTTACTAGATCAAGTCAAGTGGTCCGCCCTACATCTGCAATGTGTT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 177 CAATCAAGAAACATTTACTAGATCAAGTCAAGTGGTCCGCCCTACATCTGCAATGTGTT 236
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 CGAATTAATTAACATGAGAGCTCTATCATGATCAGGAGATGCTCCCTGATGATGATGCC 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 237 CGAATTAATTAACATGAGAGCTCTATCATGATCAGGAGATGCTCCCTGATGATGATGCC 296
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 AAGGCTTTGGGCTGCTGATCTTCTGCTGATGAGGATGATGATGATGATGATGATGATGAT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 297 AAGGCTTTGGGCTGCTGATCTTCTGCTGATGAGGATGATGATGATGATGATGATGATGAT 356
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 ATACACAGAGCCCTTGAAGACACAGGTTGAGACGAGCTGAGAAAAAATGTTCTGTG 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 357 ATACACAGAGCCCTTGAAGACACAGGTTGAGACGAGCTGAGAAAAAATGTTCTGTG 416
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 ATGAGATGATCTTCAAGAGATCATCCCGACGACCACTGCTGCTGACGAGAACAT 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 417 ATGAGATGATCTTCAAGAGATCATCCCGACGACCACTGCTGCTGACGAGAACAT 476
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 GGGGATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 477 GGGGATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 536
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 CTCGGGCTTTTGCATTTCTCTGAGCCTTTTCAGGACAGTAGATGATGATGATGATGATGATGAT 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 537 CTCGGGCTTTTGCATTTCTCTGAGCCTTTTCAGGACAGTAGATGATGATGATGATGATGATGAT 595
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 481 CGATATGATTTACTGATGATG 500
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 596 CGATATGATTTACTGATGATG 615

RESULT 15

B6333881 936 bp mRNA linear EST 27-FEB-2001
 LOCUS B6333881
 DEFINITION 602460213F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4576995 5',
 mRNA sequence.

ACCESSION B6333881
 VERSION B6333881.1 GI:13140319
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M1290 row: j column: 04
 High quality sequence stop: 598.
 Location/Qualifiers

FEATURES

1. 936
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4576995"
 /clone_id="NIH_MGC_20"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:4576995"  
/clone_1b="NIR_MSC-20"  
/issue_type="melanotic melanoma"  
/lab_host="PH108 (phage-resistant)"  
/note="organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT      237 a      235 c      249 g      215 t  
ORIGIN
```

Query Match 97.0% Score 485; DB 10; Length 936;

Best Local Similarity 99.8% Pred. No. 1,2e-139; Mismatches 0; Indels 1; Gaps 1;

Matches 496; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
OY 1 GAATTCCTGACTGCCACAGGTGTACAGAAATTTGCTTTGTTGCTGGAAGCTCT 60  
DB 51 GAATTCCTGACTGCCACAGGTGTACAGAAATTTGCTTTGTTGCTGGAAGCTCT 110  
OY 61 CAATCAAGAGACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCTCATGTGTT 120  
DB 111 CAATCAAGAGACATTTACTGAAGTCAAGTGTGCGCCCTACATCTCTCATGTGTT 170  
OY 121 CGAATATTTACATCAGAGCTCTATCGATCAGTGGAGATGTCTCCGATGTTGATGCC 180  
DB 171 CGAATATTTACATCAGAGCTCTATCGATCAGTGGAGATGTCTCCGATGTTGATGCC 230  
OY 181 AAGCTTTGGTGGCTGCTGACTTCTTCTGGTGTATGGGATGTCATCTCAACATCAAT 240  
DB 231 AAGCTTTGGTGGCTGCTGACTTCTTCTGGTGTATGGGATGTCATCTCAACATCAAT 290  
OY 241 ATCACAGAGCCCTTGAGAGAACACAGGTTGAGACGGAAGCTAGAAAAAATGTTCTGTG 300  
DB 291 ATCACAGAGCCCTTGAGAGAACACAGGTTGAGACGGAAGCTAGAAAAAATGTTCTGTG 350  
OY 301 ATGACGATGATCTTCAAGAGATCCTCCACAGCCACCAACATCGTTGCCACGAGACAT 360  
DB 351 ATGACGATGATCTTCAAGAGATCCTCCACAGCCACCAACATCGTTGCCACGAGACAT 410  
OY 361 GTGGTAGTGGCTGTGATGTATACCAACAAGGTTCTCCATTTTCAAGAAGCCAGGGT 420  
DB 411 GTGGTAGTGGCTGTGATGTATACCAACAAGGTTCTCCATTTTCAAGAAGCCAGGGT 470  
OY 421 CTCGGGCGTTTTCATTTCTCTGAGCCTGTTTCA-GGGCAGTATGATGAGTGAGGT 479  
DB 471 CTCGGGCGTTTTCATTTCTCTGAGCCTGTTTCA-GGGCAGTATGATGAGTGAGGT 530  
OY 480 TCGATATGATTTACTGG 496  
DB 531 TCGATATGATTTACTGG 547
```

Search completed: October 10, 2002, 20:27:49
Job time : 702.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 84.9333 Seconds
(without alignments)
10107.424 Million cell updates/sec

```

Title:      US-09-489-101A-9_COPY_1_500
Perfect score: 500
Sequence:   1 gaattcctgactgccacag.....cgatatgatttactgattg 500

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Maximum DB seq length: 20000000000

Pre-processing:	Minimum Match	Maximum Match
Post-processing:	Minimum Match 0%	Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802: *

1	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1980.DAT *
2	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1981.DAT *
3	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1982.DAT *
4	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1983.DAT *
5	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1984.DAT *
6	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1985.DAT *
7	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1986.DAT *
8	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1987.DAT *
9	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1988.DAT *
10	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1989.DAT *
11	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1990.DAT *
12	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1991.DAT *
13	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1992.DAT *
14	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1993.DAT *
15	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1994.DAT *
16	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1995.DAT *
17	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1996.DAT *
18	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1997.DAT *
19	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1998.DAT *
20	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA1999.DAT *
21	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA2000.DAT *
22	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA2001.DAT *
23	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA2002.DAT *
24	/SIDSI/gcgagata/hold-geneseq/genseqn-emb1/NA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	500	100.0	2281	22	AA011116	Human small cell 1
2	500	100.0	2297	22	AA158623	Human polynucleoti
3	498.4	99.7	2368	22	AA160409	Human polynucleoti
C	262.2	52.4	2898	22	ABA08061	Human ovarian and
5	262.2	52.4	2898	22	ABA08062	Human ovarian and
C	262.2	52.4	2898	22	AA06787	Human reproductive
C	7.252.2	52.4	2898	22	AA106788	Human reproductive
8	38.8	7.8	2010	23	AB106741	Drosophila melanog
9	38.8	7.8	4272	23	AB106740	Drosophila melanog

C	10	38.8	7.8	6162	23	A16J06328	Drosophila melanog
C	11	33.6	6.7	16635	22	AAFE26953	Human cancer assoc
C	12	32.6	6.5	18301	22	AA1F96894	Human excretory re
C	13	32.6	6.5	18301	22	AA1F96894	Human kidney relat
C	14	32.4	6.5	996	22	AAFE26310	Pseudomonas sp non
C	15	32.2	6.4	2187	22	AAAB83052	Human transcritp
C	16	32.2	6.4	30078	21	AAAB81520	N. meningitidis pa
C	17	32.2	6.4	349980	21	AAAF26508	Escherichia mening
C	18	32	6.4	12126	23	ABLI0434	Drosophila melanog
C	19	31.8	6.4	5151	22	ABAI16681	Human nervous syst
C	20	31.8	6.4	5151	22	AAJ050800	Human reproductive
C	21	31.4	6.3	1574	23	ABJ04959	Drosophila melanog
C	22	31.4	6.3	3748	23	ABJ04959	Drosophila melanog
C	23	31.4	6.3	3813	23	ABLI17884	Drosophila melanog
C	24	30.6	6.1	769	22	AAAI96719	Human neuroblastom
C	25	30.4	6.1	555	22	AAHI10947	Human CDNA clone (
C	26	30.2	6.0	1663	24	AAAS97010	Arabidopsis DNA for
C	27	30.2	6.0	12138	23	ABLI15540	Drosophila melanog
C	28	30.2	6.0	12161	23	ABLI15542	Drosophila melanog
C	29	30.2	6.0	32476	22	AAAK85314	Human immunohaema
C	30	30	6.0	710	22	AAI093376	Human breast cancer
C	31	30	6.0	735	22	AAAL17265	Human breast cancer
C	32	30	6.0	1422	22	AAAB89030	Escherichia coli p
C	33	30	6.0	3546	20	AAAX90855	Human peroxisome f
C	34	30	6.0	3571	24	AAAO3787	CDNA clone PL776.6
C	35	30	6.0	10899	22	ABAI15344	Human nervous syst
C	36	30	6.0	12758	22	ABAI15345	Human nervous syst
C	37	30	6.0	16066	22	AAAB89022	Escherichia coli p
C	38	29.8	6.0	842	22	AAI939960	Human neuroblastom
C	39	29.8	6.0	1541	16	AAAT04133	Helicobacter pylori
C	40	29.8	6.0	3590	18	AAAF53551	Interleukin-1 rece
C	41	29.8	6.0	4389	23	ABJ264429	Drosophila melanog
C	42	29.6	6.0	7937	21	ABJ264428	Drosophila melanog
C	43	29.6	5.9	1583	21	AAAC78952	Human secreted pte
C	44	29.6	5.9	1677	22	AAI65735	Human retinoblasto
C	45	29.6	5.9	2866	23	AAAS92572	DNA encoding novel

ALIGNMENTS

XX	RESULT 1
AC	AAD11116
ID	AAD11116 standard; DNA; 2291 BP.
AC	AAD11116;
DT	24-SEP-2001 (first entry)
XX	
DE	Human small cell lung cancer associated gene, eIF2B.
XX	
KW	Human; small cell lung cancer; therapy; hCAP; nucleic acid; NR
KW	melanoma; cancer; colon; breast; head; neck; transitional cancer
KM	leukosarcoma; eukaryotic translation initiation factor; eIF2B
XX	Novial sarcoma; cytosstatic; ds.
OS	Homo sapiens.
PX	WO200153349-A2.
PN	
XX	
PD	26-JUL-2001.
PF	19-JAN-2001; 2001WO-USO2015.
XX	
XX	21-JAN-2000; 2000US-0489101.
PA	(LUDW-) LUDWIG INST CANCER RES.
PA	(SLOK) SLOAN KETTERING INST CANCER RES.
XX	(CORR) CORNELL RES FOUND INC.
PI	Stocker E, Scalan MJ, Jager D, Old LJ, Gure AO, Chen Y;
RR	WPI; 2001-457597/49.

XX Isolated polypeptide, used to treat or prognose a disorder
 PT Characterized by expression of a hCAP e.g. cancer, is encoded by an
 PT Isolated nucleic acid comprising a hNA Group 3 or 4 molecule -
 XX
 XX Claim 57: Page 95-96; 152pp; English.
 CC The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterised by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterised by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterised by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are small and non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
 CC The present sequence is a DNA encoding human eukaryotic translation
 CC initiation factor eIF2B. This small cell lung cancer associated
 CC gene is designated as NY-SCLC-7.
 XX
 SO Sequence 2291 BP; 588 A; 540 C; 630 G; 533 T; 0 other:
 Query Match 100.0%; Score 500; DB 22; Length 2291;
 Best Local Similarity 100.0%; Pred. No. 3.3e-159;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATTCCTGACTGCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 60
 DB 1 GAATTCCTGACTGCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 60
 QY 61 CAATCAAGAAACATTTACTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 120
 DB 61 CAATCAAGAAACATTTACTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 120
 QY 121 CGAATTAATTAACATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 180
 DB 121 CGAATTAATTAACATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 180
 QY 181 AAGGCTTGGTGGCTCTGACTTCTTCTGCTGTATGGGATGTCATCTCAACATCAAT 240
 DB 181 AAGGCTTGGTGGCTCTGACTTCTTCTGCTGTATGGGATGTCATCTCAACATCAAT 240
 QY 241 ATACACAGAGCCCTTGAGAACAGAGTGTAGAGAGAGTGTAGAGAGAGTGTAGAGAG 300
 DB 241 ATACACAGAGCCCTTGAGAACAGAGTGTAGAGAGAGTGTAGAGAGAGTGTAGAGAG 300
 QY 301 ATACAGATGATCTTCAAGAGATCAATCCCAAGTCAATCTGCTGCAAGAGACAT 360
 DB 301 ATACAGATGATCTTCAAGAGATCAATCCCAAGTCAATCTGCTGCAAGAGACAT 360
 QY 361 GTGGTAGTGGCTGTGTGATACCAACAAACAGAGTTCATTTTCAGAAAGCCAGAGT 420
 DB 361 GTGGTAGTGGCTGTGTGATACCAACAAACAGAGTTCATTTTCAGAAAGCCAGAGT 420
 QY 421 CTCGGCGCTTTTTCAGATTTCTCTGAGAGCTGTTTCAGAGAGAGTGTAGAGAGT 480
 DB 421 CTCGGCGCTTTTTCAGATTTCTCTGAGAGCTGTTTCAGAGAGAGTGTAGAGAGT 480
 QY 481 CGAATGATTTACTGATTC 500
 DB 481 CGAATGATTTACTGATTC 500

DE Human polynucleotide SEQ ID NO 826.
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 EN WO200153312-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 26-DEC-2000; 2000WO-US34263.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0552317.
 XX 09-JUL-2000; 2000US-0598042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.
 XX 14-SEP-2000; 2000US-0662191.
 XX 19-OCT-2000; 2000US-0693036.
 XX 29-NOV-2000; 2000US-0727344.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao Q, Zhou P, Goodrich R, Drmanac R;
 XX
 XX WPI: 2001-442253/47.
 XX P-PsDB; AAM39467.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS
 PS Claim 1; SEQ ID NO 826; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SO Sequence 2907 BP; 734 A; 703 C; 810 G; 653 T; 7 other:
 Query Match 100.0%; Score 500; DB 22; Length 2907;
 Best Local Similarity 100.0%; Pred. No. 3.7e-159;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATTCCTGACTGCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 60
 DB 599 GAATTCCTGACTGCCACAGGTGTACAGAAACATTTGCTTTGCTGGAAGCTGCT 658
 QY 61 CAATCAAGAAACATTTACTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 120
 DB 659 CAATCAAGAAACATTTACTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 718
 QY 121 CGAATTAATTAACATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 180

Db 719 CGAATAATTCATCAGACTCTATGCATCTGAGAGATGTCCTCCGTGATGTGATGCC 778

QY 181 AAGCTTTGGTGGCTCTGACCTTTCTTGTGTGTATGGGATGTCATCTCAACATCAAT 240

Db 779 AAGCTTTGGTGGCTCTGACCTTTCTTGTGTGTATGGGATGTCATCTCAACATCAAT 838

QY 241 ATCAGCAGACCCCTTGAGAGACACAGAGTTGAGACGAGTATGAGAAAAATGTTCTGTG 300

Db 839 ATCAGCAGACCCCTTGAGAGACACAGAGTTGAGACGAGTATGAGAAAAATGTTCTGTG 898

QY 301 ATACAGCATATCTTCAAGAGATCACCACCCACCCCACTGCTGGCCAGAGACAT 360

Db 899 ATACAGCATATCTTCAAGAGATCACCACCCACCCCACTGCTGGCCAGAGACAT 958

QY 361 GTGCTAGTGGCTGTGATAGTACCAACAAACAGAGTTCTTCATTTTCAGAGACCCAGGTT 420

Db 959 GTGCTAGTGGCTGTGATAGTACCAACAAACAGAGTTCTTCATTTTCAGAGACCCAGGTT 1018

QY 421 CTCGCGGCTTTGATTTCTCTGAGCCCTGTTTCAGGCGATGATGATGAGAGTT 480

Db 1019 CTCGCGGCTTTGATTTCTCTGAGCCCTGTTTCAGGCGATGATGATGAGAGTT 1078

QY 481 CGATATGATTTACTGATTTG 500

Db 1079 CGATATGATTTACTGATTTG 1098

RESULT 3
AA160409 standard; cDNA: 2968 BP.

AC AA160409:

DF 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4398.

DE Human polynucleotide SEQ ID NO 4398.

XX Human; nootropic; immunosuppressant; cytoskeletal; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokine; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

OS Homo sapiens.

XX MO20015312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QH, Zhou P, Goodrich R, Dimauc RT;

XX WPI: 2001-442253/47.

XX P-PSDB: AAM1253.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX

PS Claim 1; SEQ ID NO 4398; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA44213) with nootropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 2968 BP; 728 A; 719 C; 825 G; 696 T; 0 other:

Query Match 99.7%; Score 498.4; DB 22; Length 2968;

Matches 499; Conservatively 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAATTCCTGACTGCGACAGGTGTACAGAGAAACATTTGCTTTGTTGCTGAGAACTGCT 60

Db 262 GAATTCCTGACTGCGACAGGTGTACAGAGAAACATTTGCTTTGTTGCTGAGAACTGCT 321

QY 61 CAATCAAGAAACATTTAGTGAAGTCAAAAGTGGTCCGCTTACATCTCAATGAGTT 120

Db 322 CAATCAAGAAACATTTAGTGAAGTCAAAAGTGGTCCGCTTACATCTCAATGAGTT 381

QY 121 CGAATAATTCATCAGACTCTATGCATCTGAGAGATGTCATCTCAACATCAAT 180

Db 382 CGAATAATTCATCAGACTCTATGCATCTGAGAGATGTCATCTCAACATCAAT 441

QY 181 AAGCTTTGGTGGCTCTGACCTTTCTTGTGTGTATGGGATGTCATCTCAACATCAAT 240

Db 442 AAGCTTTGGTGGCTCTGACCTTTCTTGTGTGTATGGGATGTCATCTCAACATCAAT 501

QY 241 ATCAGCAGACCCCTTGAGAGACACAGAGTTGAGACGAGTATGAGAAAAATGTTCTGTG 300

Db 502 ATCAGCAGACCCCTTGAGAGACACAGAGTTGAGACGAGTATGAGAAAAATGTTCTGTG 561

QY 301 ATACAGCATATCTTCAAGAGATCACCACCCACCCCACTGCTGGCCAGAGACAT 360

Db 562 ATACAGCATATCTTCAAGAGATCACCACCCACCCCACTGCTGGCCAGAGACAT 621

QY 361 GTGCTAGTGGCTGTGATAGTACCAACAAACAGAGTTCTTCATTTTCAGAGACCCAGGTT 420

Db 622 GTGCTAGTGGCTGTGATAGTACCAACAAACAGAGTTCTTCATTTTCAGAGACCCAGGTT 681

QY 421 CTCGCGGCTTTGATTTCTCTGAGCCCTGTTTCAGGCGATGATGATGAGAGTT 480

Db 682 CTCGCGGCTTTGATTTCTCTGAGCCCTGTTTCAGGCGATGATGATGAGAGTT 741

QY 481 CGATATGATTTACTGATTTG 500

Db 742 CGATATGATTTACTGATTTG 761

RESULT 4
ABA08061/C
ID ABA08061 standard; DNA: 2898 BP.

XX ABA08061:

XX 11-JAN-2002 (first entry)

XX Human ovarian and breast cancer associated polynucleotide SEQ ID NO 856.
DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX

PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-488786/53.
 XX
 PR New isolated ovarian and/or breast cancer related nucleic acids and
 PT polypeptides, useful for diagnosing, treating and/or preventing human
 PT diseases and disorders, particularly ovarian and/or breast cancer -
 XX
 PS Disclosure: SEQ ID NO 856; 577bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA07454-ABA08224) and proteins
 CC (ABA0743-ABA0980) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful.
 CC In the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPRO at ftp.wipro.int/pub/published_pct_sequences.
 CC
 XX
 SO Sequence 2898 BP; 796 A; 744 C; 600 G; 758 T; 0 other;
 Query Match 52.4%; Score 262.2; DB 22; Length 2898;
 Best Local Similarity 79.4%; Pred. No. 2.8e-78;
 Matches 366; Conservative 0; Mismatches 3; Indels 92; Gaps 1;

RESULT 5
 ABA08062/c
 ID ABA08062 standard; DNA; 2898 BP.
 XX
 AC ABA08062;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human ovarian and breast cancer associated polynucleotide SEQ ID NO 857.
 XX
 KW Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KW vlnnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200155325-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01345.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225470.
 PR 14-AUG-2000; 2000US-0225477.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241285.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0246174.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 03-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251989.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI: 2001-488766/53.
 XX
 XX New isolated ovarian and/or breast cancer related nucleic acids and
 PT polypeptides, useful for diagnosing, treating and/or preventing human
 PT diseases and disorders, particularly ovarian and/or breast cancer -
 XX
 PS Disclosure; SEQ ID NO 857; 577bp + sequence listing; English.
 XX
 CC The invention relates to novel genes (ABA07454-ABA08224) and proteins
 CC (ABA07454-ABA0980) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 2898 BP; 796 A; 744 C; 600 G; 758 T; 0 other;
 Query Match 52.4%; Score 262.2; DB 22; Length 2898;
 Best Local Similarity 79.4%; Pred. No. 2.8e-78;
 Matches 366; Conservative 0; Mismatches 3; Indels 92; Gaps 1;
 QY 76 TTTACTGAGTCAAGTGGTCCGCCCTACATCTCTCAATGTTGTCGAATTAATCATCA 135
 Db 644 TTTTGAAGTCAAGTGGTCCGCCCTACATCTCTCAATGTTGTCGAATTAATCATCA 585
 QY 136 GAGCTATATGATCACTGGAGATGCTCTCGATGTTGATGTCGAAGCTTGTGTCGC 195
 Db 584 GAGCTATATGATCACTGGAGATGCTCTCGATGTTGATGTCGAAGCTTGTGTCGC 525
 QY 196 TCTGACTTCTTCTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 255
 Db 524 TCTGACTTCTTCTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 465
 QY 256 GAGGACAC----- 264
 Db 464 GAGGACACAGTCAAGATGGAATAATGACAGGATTAAGACACGACGACCC 405
 QY 265 -----AGGTTGAGACGAGACTAG 283
 Db 404 TGAGACTGTTTCTTTCAGTCTGTCCTCTGTCCTTATAGTGTAGACGAGACTAG 345

```
OY 284 AAAAAAAAAATGTTCTGTGATGACGATGATCTTCAAGAGTCATCCCGACCCCACTC 343
DB 344 AAAAAAAAAATGTTCTGTGATGACGATGATCTTCAAGAGTCATCCCGACCCCACTC 285
OY 344 GTTGCACGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
DB 284 GTTGCACGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 225
OY 404 TTGAGAGACCCGAGGCTCCGCGGTTTGATTCCTCTG 444
DB 224 TTGAGAGACCCGAGGCTCCGCGGTTTGATTCCTCTG 184

RESULT 6
AAL06787/c
ID AAL06787 standard; DNA; 2898 BP.
AC AAL06787;
XX 21-NOV-2001 (first entry)
DT
XX Human reproductive system related antigen DNA SEQ ID NO: 9475.
XX Human reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; de.
XX Homo sapiens.
XX MO200155320-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01339.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205155.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216890.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225277.
XX 14-AUG-2000; 2000US-0225278.
XX 14-AUG-2000; 2000US-0225279.
XX 18-AUG-2000; 2000US-0225279.
XX 22-AUG-2000; 2000US-0225681.
XX 22-AUG-2000; 2000US-0225681.
XX 22-AUG-2000; 2000US-0225681.
XX 23-AUG-2000; 2000US-0225789.
XX 30-AUG-2000; 2000US-022824.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 08-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0233397.
XX 14-SEP-2000; 2000US-0233398.
XX 14-SEP-2000; 2000US-0233399.
XX 14-SEP-2000; 2000US-0233400.
XX 14-SEP-2000; 2000US-0233401.
XX 14-SEP-2000; 2000US-0233403.
XX 14-SEP-2000; 2000US-0233404.
XX 14-SEP-2000; 2000US-0233405.
XX 21-SEP-2000; 2000US-0234273.
XX 21-SEP-2000; 2000US-0234274.
XX 25-SEP-2000; 2000US-0234997.
XX 25-SEP-2000; 2000US-0234998.
XX 26-SEP-2000; 2000US-0235484.
XX 27-SEP-2000; 2000US-0235834.
XX 27-SEP-2000; 2000US-0235836.
XX 29-SEP-2000; 2000US-0236327.
XX 29-SEP-2000; 2000US-0236367.
XX 29-SEP-2000; 2000US-0236368.
XX 29-SEP-2000; 2000US-0236369.
XX 29-SEP-2000; 2000US-0236370.
XX 02-OCT-2000; 2000US-0236802.
XX 02-OCT-2000; 2000US-0237037.
XX 02-OCT-2000; 2000US-0237038.
XX 02-OCT-2000; 2000US-0237039.
XX 02-OCT-2000; 2000US-0237040.
XX 13-OCT-2000; 2000US-0239335.
XX 13-OCT-2000; 2000US-0239337.
XX 13-OCT-2000; 2000US-0239337.
XX 20-OCT-2000; 2000US-0241221.
XX 20-OCT-2000; 2000US-0241785.
XX 20-OCT-2000; 2000US-0241786.
XX 20-OCT-2000; 2000US-0241787.
XX 20-OCT-2000; 2000US-0241808.
XX 20-OCT-2000; 2000US-0241809.
XX 20-OCT-2000; 2000US-0244617.
XX 01-NOV-2000; 2000US-0244674.
XX 08-NOV-2000; 2000US-0246475.
XX 08-NOV-2000; 2000US-0246476.
XX 08-NOV-2000; 2000US-0246477.
XX 08-NOV-2000; 2000US-0246478.
XX 08-NOV-2000; 2000US-0246524.
XX 08-NOV-2000; 2000US-0246524.
XX 08-NOV-2000; 2000US-0246525.
XX 08-NOV-2000; 2000US-0246526.
XX 08-NOV-2000; 2000US-0246527.
XX 08-NOV-2000; 2000US-0246528.
XX 08-NOV-2000; 2000US-0246529.
XX 08-NOV-2000; 2000US-0246610.
XX 08-NOV-2000; 2000US-0246611.
XX 08-NOV-2000; 2000US-0246613.
XX 17-NOV-2000; 2000US-0249207.
XX 17-NOV-2000; 2000US-0249208.
XX 17-NOV-2000; 2000US-0249209.
XX 17-NOV-2000; 2000US-0249210.
XX 17-NOV-2000; 2000US-0249211.
XX 17-NOV-2000; 2000US-0249212.
XX 17-NOV-2000; 2000US-0249213.
XX 17-NOV-2000; 2000US-0249214.
```

Query Match	Best Local Similarity	Score	262.2;	DB 22;	Length 2896;
Matches	366;	Conservative	0;	Mismatches 3;	Indels 92; Gaps 1.
50	Sequence 2898 BP; 796 A; 744 C; 600 G; 758 T; 0 other;				
51	Query Match	52.4%;	Score 262.2;	DB 22;	Length 2896;
52	Best Local Similarity	79.4%;	Pred. No. 2.8e-78;		
53	Matches	366;	Conservative	0;	Mismatches 3; Indels 92; Gaps 1.
54	76	TTTACTGAGTCAAAAGTGTGCGCCGCCCTACATCTCTCAATGTGTTGCAATTAATTACATCA	135		
55	11				
56	644	TTTGGGAAGTCAAAAGTGTGCGCCGCCCTACATCTCTCAATGTGTTGCAATTAATTACATCA	585		
57	136	GAGCTCATATGACATCAGTGGAGAGATGTGCTCCGCTGATGTTGATAGCCCAAGGCTTTGGTACGA	195		
58	11				
59	584	GAGCTCATATGACATCAGTGGAGAGATGTGCTCCGCTGATGTTGATAGCCCAAGGCTTTGGTACGA	525		
60	196	TCTACTTTTCTGTTGGTGTGATGGAGTATCATCTCAATCAATCATATATCACAGAGCCCTT	255		
61	524	TCTACTTTTCTGTTGGTGTGATGGAGTATCATCTCAATCAATCATATATCACAGAGCCCTT	465		
62	256	GAGGAACAC-----	264		
63	11				
64	464	GAGGAACACAGTGTGATGATGGAGAAATATACACAGAACAGGTTAAACACACAGAGGCC	405		
65	265	-----			
66	404	TGAGACATGCTTTTTCAGAGTGTCTGTCGCCCTGCTCTTATAGGTTGAGAGAGAGTAA	283		
67	284	AAAAAATATTTTCTGATGATGAGATGATCTTTCAGAGATCATCCCCAGCCACCCAACTC	343		
68	344	AAAAAATATTTTCTGATGATGAGATGATCTTTCAGAGATCATCCCCAGCCACCCAACTC	285		
69	344	GTTCACGAGAACAAATGTGTAGTGGCTGTGATGATGATACCAACAAAGGTTCTCATTT	403		

RESULT 7
AA106788/G
ID AA106788 standard; DNA; 2898 BP.
XX AA106788;
XX DT 21-NOV-2001 (first entry)
XX DE Human reproductive system related antigen DNA SEQ ID NO: 9476.
XX RM Human; reproductive system related antigen; reproductive system disorder.
XX KW cancer; gene therapy; ds.
XX OS Homo sapiens.
XX MO200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US013359.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186360.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209457.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 11-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226686.
XX PR 23-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227609.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 03-SEP-2000; 2000US-0229345.
XX PR 03-SEP-2000; 2000US-0229509.
XX PR 06-SEP-2000; 2000US-0230457.
XX PR 06-SEP-2000; 2000US-0230458.

Db 224 TTCAGAACCCAGGCTCCGGCTTTGCATTCTCTG 184

RESULT 8
ABL06741
ID ABL06741 standard; cDNA; 2010 BP.

XX ABL06741;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 14705.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX P-PSDB; ABB62638.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 14705; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB101840-AB116175), expressed DNA

XX sequences (AB101840-AB116175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2010 BP; 463 A; 548 C; 596 G; 403 T; 0 other;

XX

ID ABL06740 standard; cDNA; 4272 BP.

XX ABL06740;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 14702.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX P-PSDB; ABB62637.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 14702; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB101840-AB116175), expressed DNA

XX sequences (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 4272 BP; 1069 A; 1157 C; 1092 G; 954 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 13466.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 OS
 PN MO20017042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW.
 XX WPI, 2001-656860/75.
 DR P-PSDB; ABB62225.
 XX
 PS New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PI interactions -
 XX
 PS Claim 1; SEQ ID NO 13466; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 6162 BP; 1591 A; 1658 C; 1564 G; 1349 T; 0 other:
 Query Match 7.8%; Score 38.8; DB 23; Length 6162;
 Best Local Similarity 51.8%; Pred. No. 0.041;
 Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
 QY 118 GTTCGATATTTACATCAGACTCTATGATCGACTGGAGATGTCCTCCGATGTTGAT 177
 DB 422 GTCCAGTGTATGCGCGGAGGAGGTTGCTGCTGGCGCATGCGATGCGATTGGAC 363
 QY 178 GCCAAGGCTTTGGTGGCTCTGCTGCTGCTGATGATGAGATCTCAACATC 237
 DB 362 AATAGGCGGATGATCCGATGCTGCTGCTGCGCGCATGAGGATGAGCAATGCC 303
 QY 238 AATATCACAGAGCCCTTGAAGACAGAGTTGAGACAGAGAGAGAGAA 287
 DB 302 GATCTGGCGCCAGTGTGAGACAGCAACAGCGGCGCAAGTTGACAA 253
 RESULT 11
 AAF26953/c
 ID AAF26953 standard; cDNA; 3635 BP.
 XX
 AC AAF26953;
 XX
 DT 10-APR-2001 (first entry)
 XX
 DE Human cancer associated antigen precursor HOM-TES-84/6 cDNA SEQ ID NO:5.
 XX
 KW Human; cancer associated antigen precursor; cancer associated antigen;
 KW seminoma; HLA; human leukocyte antigen; cytosolic; gene therapy;

KW vaccine; cancer; ss.
 OS Homo sapiens.
 XX
 PN MO200100874-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-US17207.
 XX
 PR 30-JUN-1999; 99US-0346498.
 XX
 PA (LUDWIG-) LUDWIG INST CANCER RES.
 XX
 PI Sahlin U, Tureci O, Pfeundschn M.
 XX WPI: 2001-112465/12.
 DR P-PSDB; AAB75607.
 XX
 PS Diagnosing a disorder characterized by expression of a human cancer
 PT associated antigen precursor, comprises detecting interaction of an
 PI agent with a nucleic acid molecule encoding the antigen precursor -
 XX
 PS Claim 48; Page 90-91; 126pp; English.
 XX
 CC The present invention describes a method for diagnosing a disorder
 CC characterised by expression of a human cancer associated antigen (CAA)
 CC precursor (I) coded by a NA Group 1 nucleic acid molecule (NI)
 CC comprising contacting the biological sample with an agent (A) that
 CC specifically binds to NI, (I) or its fragment, complexed with an human
 CC leukocyte antigen (HLA) molecule and determining the interaction between
 CC the agent and NI or (I). (I) has cytosolic activity and can be used in
 CC gene therapy and vaccine production. The method can be used for treating
 CC a subject with a condition characterised by expression of (I) in cells
 CC of a subject. The present sequence represents a human cancer associated
 CC antigen precursor cDNA sequence which is used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 3635 BP; 1162 A; 741 C; 797 G; 935 T; 0 other:
 Query Match 6.7%; Score 33.6; DB 22; Length 3635;
 Best Local Similarity 55.0%; Pred. No. 1.8;
 Matches 66; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 21 TGTACGGAACATTTGCTTTGCTGTGGAAGCTGTCATCAAGACATTTACT 80
 DB 559 TGCTATGAGACACACTTCTATGTGTGTGCTGACGAGAAATCTTATTAATTCCA 500
 QY 81 GAAGTCAAGTGTGCTGCGCCCTACATCTCTCAATGTGCTGATTAATTCATGACACT 140
 DB 499 GATTTGAAAGTACACACAGTTTATTTCTTTGCTTTGCTGCTTTCAATGAGAGCT 440
 RESULT 12
 AAI98944/c
 ID AAI98944 standard; DNA; 18301 BP.
 XX
 AC AAI98944;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Human excretory related polynucleotide SEQ ID NO 708.
 XX
 DE Human; excretory; neuroprotective; cytosolic; dermatological; vituicide;
 KW immunosuppressive; anti-inflammation; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antisticking; antianemic; antiarthritic; cancer;
 KW antipneumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiparasitic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; antidiabetic; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
 KW excretory system; ds.
 XX
 OS Homo sapiens.

```
XX PN WO200155313-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01323.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216860.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236337.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241231.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0246417.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249220.
PR 17-NOV-2000; 2000US-0249221.
PR 17-NOV-2000; 2000US-0249222.
PR 17-NOV-2000; 2000US-0249223.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249225.
PR 17-NOV-2000; 2000US-0249226.
PR 17-NOV-2000; 2000US-0249227.
PR 17-NOV-2000; 2000US-0249228.
PR 17-NOV-2000; 2000US-0249229.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA ) HUMAN GENOME SCI INC.
XX PA
XX
```

PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-465569/50.
DR

PT Isolated nucleic acid molecule encoding excretory system antigen is
PR used in preventing, treating or ameliorating a medical condition -
XX
PS Example 2; SEQ ID NO 708; 574bp + Sequence Listing; English.

CC The invention relates to novel excretory system related human
CC polynucleotides (AA198567-AA199503) and the encoded proteins
CC (AA199594-AA199613) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy, especially
CC disorders related to the excretory system. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral ischaemia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 18301 BP; 5076 A; 4392 C; 3842 G; 4991 T; 0 other;

Query Match 6.54; Score 32.6; DB 22; Length 18301;
Best Local Similarity 54.64; Pred. No. 9.5; Mismatches 0; Gaps 0;
Matches 65; Conservative 0; Indels 0;

QY 133 TCAGAGCTTATCGATCGAGATGCTCCGTCGATGTCATGCCAAGCTTTGGTG 192
DB 198 TCCGGGCTTCAATAGCTTCTTGTATGATGTAATTTGATGATGATTTGTTG 139
QY 193 CGCTGCTGCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 251
DB 138 AACTGTGCTTAAAGCTTTCATTAATAAAGCTTCTTCAATCAATATCTCTGTC 80

RESULT 13
AA163294/C
ID AA163294 standard; DNA; 18301 BP.
XX

AC AA163294;
XX

D7 22-OCT-2001 (first entry)
XX

DE Human kidney related polynucleotide SEQ ID NO 609.
XX

KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatocytic; antidiabetic;
KW antiinflammatory; antitumor; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX

OS Homo sapiens.
XX

PN WO200153323-A2.
XX

PD 02-AUG-2001.
XX

PF 17-JAN-2001; 2001WO-US01343.
XX

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR

PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205513.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218289.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225277.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231415.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 18.7333 Seconds
(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101A-8_COPY_1_500

Perfect score: 500
Sequence: 1 ggcagctggagccgcctgcgcgc.....aatgtctatcgtatccagctg 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	53	10.6	7218	1	US-08-232-463-14	Sequence 14, Appl
2	51.8	10.4	3083	1	US-08-346-849-1	Sequence 1, Appl
3	51.8	10.4	3083	2	US-08-293-284A-1	Sequence 1, Appl
4	35.2	7.0	377	4	US-08-961-810-19	Sequence 19, Appl
5	35.2	7.0	377	4	US-08-352-902D-19	Sequence 1, Appl
6	34.8	7.0	1293	3	US-09-007-484-1	Sequence 1, Appl
7	34.8	7.0	1293	3	US-09-309-682-1	Sequence 1, Appl
8	34.6	6.9	1924	3	US-08-961-083-159	Sequence 159, App
9	34.6	6.9	2223	1	US-08-257-073-4	Sequence 4, Appl
10	34.4	6.9	2547	3	US-08-262-220-7	Sequence 7, Appl
11	34.4	6.9	2547	3	US-08-471-733-7	Sequence 7, Appl
12	34.4	6.9	2547	3	US-08-468-878-7	Sequence 7, Appl
13	34.4	6.9	2547	3	US-08-750-494-7	Sequence 7, Appl
14	33.2	6.6	4800	3	US-09-106-638-1	Sequence 1, Appl
15	33	6.6	1578	1	US-08-681-129-1	Sequence 1, Appl
16	32.8	6.6	30001	1	US-08-125-468-1	Sequence 1, Appl
17	32.8	6.6	30001	2	US-08-474-933-1	Sequence 1, Appl
18	32.6	6.5	600	4	US-09-228-986-62	Sequence 62, Appl
19	32.6	6.5	20137	4	US-09-262-773-206	Sequence 206, App
20	32.6	6.5	20138	4	US-09-262-773-9	Sequence 9, Appl
21	32.6	6.5	23071	4	US-09-262-773-210	Sequence 210, App
22	32.2	6.4	1789	4	US-08-936-165A-116	Sequence 116, App
23	32.2	6.4	2067	4	US-09-106-194-11	Sequence 11, Appl
24	32	6.4	2089	1	US-08-552-142A-1	Sequence 1, Appl
25	32	6.4	2089	1	US-08-910-973-1	Sequence 1, Appl
26	32	6.4	2089	5	PCT-US95-05741-1	Sequence 1, Appl
27	32	6.4	3708	1	US-08-185-232A-1	Sequence 1, Appl

28	32	6.4	3708	1	US-08-416-523-1	Sequence 1, Appl
29	32	6.4	3708	4	US-08-789-478-1	Sequence 1, Appl
30	31.6	6.3	605	4	US-09-328-111-712	Sequence 112, App
31	31.6	6.3	3211	4	US-08-574-959A-8	Sequence 8, Appl
32	31.6	6.3	3211	4	US-08-357-014-8	Sequence 8, Appl
33	31.6	6.3	3901	2	US-08-574-959A-6	Sequence 6, Appl
34	31.6	6.3	3901	4	US-09-357-014-6	Sequence 6, Appl
35	31.4	6.3	3489	2	US-08-728-327A-1	Sequence 1, Appl
36	31.4	6.3	5394	3	US-08-688-376-1	Sequence 1, Appl
37	31.4	6.3	32207	2	US-08-770-379-20	Sequence 20, Appl
38	31.4	6.3	32207	4	US-08-757-669A-20	Sequence 20, Appl
39	31.4	6.3	32207	4	US-09-230-371A-20	Sequence 20, Appl
40	31.2	6.2	2458	4	US-08-611-587-6	Sequence 6, Appl
41	31	6.2	1050	4	US-09-137-855-1	Sequence 1, Appl
42	30.8	6.2	9468	4	US-08-325-547-10	Sequence 10, Appl
43	30.6	6.1	1041	4	US-09-371-913A-10	Sequence 10, Appl
44	30.6	6.1	4766	5	PCT-US93-07261-10	Sequence 10, Appl
45	30.6	6.1	13987	2	US-08-804-227C-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: VA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F1s
US-08-232-463-14
Query Match 10.6%; Score 53; DB 1; Length 7218;

Best Local Similarity 5.4%; Pred. No. 1.1e-06;
Matches 17; Conservative 180; Mismatches 120; Indels 0; Gaps 0;

QY 135 GTGGAGAGATGTTGAGCTTTTGTAAAGAGAGAAAGAAATGCTTCTCTTTT 194
1463 GTAGTAAAGAGATAGAAATTTGTACRRRRRRRRRRRRRRRRRRRRRRRRRR 1404
QY 195 CAGGAAGCTGAGATAGAAAGATTTCCGAGATCAGAGATGAAAGATGACAGT 254
1403 RRR 1344
QY 255 GAAGAGTTCCATGCTGAAACACTGATCCAGAGCTGGAAGCAAGATCATTA 314
1343 RRR 1284
QY 315 GCAGTTCTGAGCTGCGCATGTGATAGAGCTACAGAGAGAGATCAAGAGCT 374
1283 RRR 1224
QY 375 CATAAAGCAATGTTTAAACATCACCAGAGCAAGAGAGAGCTGTGAACCA 434
1223 RRR 1164
QY 435 AAGAAGAGATATGA 451
DB 1163 RRR 1147

RESULT 2

US-08-346-849-1
Sequence 1, Application US/08346849
Patent No. 5670483

GENERAL INFORMATION:
APPLICANT: Zhang, Shuguang
APPLICANT: Lockshin, Curtis
APPLICANT: Rich, Alexander
APPLICANT: Holmes, Todd
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,849
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 1292..2590
OTHER INFORMATION: /product= "zuotin"

US-08-346-849-1

Query Match 10.4%; Score 51.8; DB 1; Length 3083;
Best Local Similarity 55.7%; Pred. No. 1.6e-06;
Matches 123; Conservative 0; Mismatches 92; Indels 6; Gaps 1;

QY 276 ACACTTGATCCCAAGACTGGAAGAACCAAGATCATTAATGCACTTGGACCTGGCCAT 335
1550 ACTCAAGATGCTAGAGATCGGAAGAACTGGCATTTGATGCTGATGCTGCTAAG 1609
QY 336 GTGAGATAGAGCTGACAGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 395
1610 TTGCGTTTCAGAGCTACTGAAATCATATCAAGCTTCACAGAAACAGTTGTCAAG 1669
QY 396 CATCACCCAGACCAAG 455
1670 TACCATCCAGACCAAG 1723
QY 456 TTGACTTGATCACTAAG 496
DB 1724 TTCAGATTAATTCAAAAGGCGCTTGAAACTTTGACTGATTC 1764

RESULT 3

US-08-293-284A-1
Sequence 1, Application US/08293284A
Patent No. 5955343

GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuguang
APPLICANT: Rich, Alexander
APPLICANT: DiPersio, C. Michael
APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,284A
FILING DATE: 22-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

[illegible]

TELEX: 248633
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Borrelia garinii
STRAIN: IP90
IMMEDIATE SOURCE:
CLONE: pJB-101
FEATURE:
NAME/KEY: CDS
LOCATION: 380..2245
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 380..442
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 443..2242
US-08-262-220-7

Query Match
Best Local Similarity 56.8%; Score 34.4; DB 3; Length 2547;
Matches 84; Conservative 0; Mismatches 61; Indels 3; Gaps 1;

QY 342 TACAGGCTACACAGACAGATCAAGCAGCTCATTAAGCAATGTTTAAACATCAC 401
DB 2472 TATAGTCTGTGATGATGATTAAGCAATGTTTAAACATCAC 2413
QY 402 CCAGACAAAGGAAAGAGAGAGTGTGACCAATTAAGAGAGATATGACTTCACT 461
DB 2412 CCTGACA--GGTTGGCAATGATCCTGTAGAGCAAAAAGATGCAATGATTTAT 2356
QY 462 TGCATTAAGCTTATGAAATGTTAT 489
DB 2355 AAATTCAGATGCTTATGAAAGATTT 2328

RESULT 11
US-08-471-733-7/C
Sequence 7, Application US/08471733
Patent No. 6068842
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
APPLICANT: BARBOUR ALAN G.
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELLIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/471.733
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/262,220
FILING DATE: 20-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BERGSTROM=3
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Borrelia garinii
STRAIN: IP90
IMMEDIATE SOURCE:
CLONE: pJB-101
FEATURE:
NAME/KEY: CDS
LOCATION: 380..2245
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 380..442
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 443..2242
US-08-471-733-7

Query Match
Best Local Similarity 56.8%; Score 34.4; DB 3; Length 2547;
Matches 84; Conservative 0; Mismatches 61; Indels 3; Gaps 1;

QY 342 TACAGGCTACACAGACAGATCAAGCAGCTCATTAAGCAATGTTTAAACATCAC 401
DB 2472 TATAGTCTGTGATGATGATTAAGCAATGTTTAAACATCAC 2413
QY 402 CCAGACAAAGGAAAGAGAGTGTGACCAATTAAGAGAGATATGACTTCACT 461
DB 2412 CCTGACA--GGTTGGCAATGATCCTGTAGAGCAAAAAGATGCAATGATTTAT 2356
QY 462 TGCATTAAGCTTATGAAATGTTAT 489
DB 2355 AAATTCAGATGCTTATGAAAGATTT 2328

RESULT 12
US-08-468-878-7/C
Sequence 7, Application US/08468878
Patent No. 6090586
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
APPLICANT: BARBOUR ALAN G.
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELLIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/468.878
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/262,220
FILING DATE: 20-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005

REFERENCE/DOCKET NUMBER: BERGSTROM-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Borrelia garinii
STRAIN: Ip90
IMMEDIATE SOURCE:
CLONE: PUB-101
FEATURE:
NAME/KEY: CDS
LOCATION: 380..2245
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 380..442
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 443..2242
US-08-468-878-7

Query Match 6.9%; Score 34.4; DB 3; Length 2547;
Best Local Similarity 56.8%; Pred. No. 0.49;
Matches 84; Conservative 0; Mismatches 61; Indels 3; Gaps 1;

QY 342 TACAGGCTACAGACAGACATCAAGCAGCTCATTAAGCATGTTTAAACATCAC 401
DB 2472 TATAGTCTGTGCTGTGATTAAGGCTTATTAAGGCTTGTATTAATATCAT 2413
QY 402 CCAGACAAAGGAAAGCAGCTGTGAACCAATTAAGAGAGATATGACTTCTACT 461
DB 2412 CCGACAC---GGTTGCAATGATCTGTAAAGCAAAAGATGCAATGATTAATTAT 2356
QY 462 TGCATACCTAAAGCTTATGAATGTTAT 489
DB 2355 AAAATTCAAGATGCTTATGAAAAGATT 2328

RESULT 13
US-08-750-494-7/c
Sequence 7, Application US/08750494
Patent No. 6204018
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
APPLICANT: BARBOUR ALAN G.
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,494
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/262,220
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BERGSTROM-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Borrelia garinii
STRAIN: Ip90
IMMEDIATE SOURCE:
CLONE: PUB-101
FEATURE:
NAME/KEY: CDS
LOCATION: 380..2245
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 380..442
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 443..2242
US-08-750-494-7

Query Match 6.9%; Score 34.4; DB 4; Length 2547;
Best Local Similarity 56.8%; Pred. No. 0.49;
Matches 84; Conservative 0; Mismatches 61; Indels 3; Gaps 1;

QY 342 TACAGGCTACAGACAGACATCAAGCAGCTCATTAAGCATGTTTAAACATCAC 401
DB 2472 TATAGTCTGTGCTGTGATTAAGGCTTATTAAGGCTTGTATTAATATCAT 2413
QY 402 CCAGACAAAGGAAAGCAGCTGTGAACCAATTAAGAGAGATATGACTTCTACT 461
DB 2412 CCGACAC---GGTTGCAATGATCTGTAAAGCAAAAGATGCAATGATTAATTAT 2356
QY 462 TGCATACCTAAAGCTTATGAATGTTAT 489
DB 2355 AAAATTCAAGATGCTTATGAAAAGATT 2328

RESULT 14
US-09-106-638-1
Sequence 1, Application US/09106638
Patent No. 6093556
GENERAL INFORMATION:
APPLICANT: KANJI NAKAMURA
APPLICANT: HIROAKI ISHIDA
TITLE OF INVENTION: GENE RECOMBINANT WITH BIODEGRADABILITY
TITLE OF INVENTION: FOR CHLORINATED ETHYLENE AND BIO-TREATMENT OF
TITLE OF INVENTION: CHLORINATED ETHYLENE THEREWITH
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: C/O KURITA WATER INDUSTRIES LTD.
STREET: 4-7, Nishi-Shinjuku 3-Chome
CITY: Shinjuku-Ku
STATE: Tokyo
COUNTRY: Japan
ZIP: 160-0023
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.5 inches, 1.44MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/106,638
FILING DATE:

ORIGIN

Query Match 100.0%; Score 500; DB 10; Length 761;
 Best Local Similarity 100.0%; Pred. No. 4.6e-117;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAGCTGAGCCGCTGCGCCGCAACCGGGCTAGACCCGCGCCATCATGCTGCTTCCCAA 60
 DB 69 GGGAGCTGAGCCGCTGCGCCGCAACCGGGCTAGACCCGCGCCATCATGCTGCTTCCCAA 128
 QY 61 GGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 DB 129 GGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 188
 QY 121 GTCAAGTTGAACCTGTGGAGAGATGGTTTGAAGCTTTTGTAAAGAGAGAAACAGAAATG 180
 DB 189 GTCAAGTTGAACCTGTGGAGAGATGGTTTGAAGCTTTTGTAAAGAGAGAAACAGAAATG 248
 QY 181 GTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAAGATTAAGAAAGATTAAGAAAGATG 240
 DB 249 GTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAAGATTAAGAAAGATTAAGAAAGATG 308
 QY 241 AAGAAATTCAGTTGGAAGATTTCCCATGCTGAAACACTTGAATCCCAAGACTGGAGA 300
 DB 309 AAGAAATTCAGTTGGAAGATTTCCCATGCTGAAACACTTGAATCCCAAGACTGGAGA 368
 QY 301 ACCAAGATCATATGCAAGTTTGGAGCTTGGCATGTGAGATCAAGGCTACACAGAGAC 360
 DB 369 ACCAAGATCATATGCAAGTTTGGAGCTTGGCATGTGAGATCAAGGCTACACAGAGAC 428
 QY 361 AGATCAAGAGCTCATTAAGCAATGTTTAAACATCAACCCAGAACGAGAAAGCAG 420
 DB 429 AGATCAAGAGCTCATTAAGCAATGTTTAAACATCAACCCAGAACGAGAAAGCAG 488
 QY 421 CTGGTGACCAATTAAGAAAGAGATTAAGTACTTCACTTGCATTAACATTAAGCTTATG 480
 DB 489 CTGGTGACCAATTAAGAAAGAGATTAAGTACTTCACTTGCATTAACATTAAGCTTATG 548
 QY 481 AATGTTATCTGATCCAGTG 500
 DB 549 AATGTTATCTGATCCAGTG 568

RESULT 2

AU121047 844 bp mRNA linear EST 19-OCT-2000
 LOCUS AU121047 HEMBI Homo sapiens cDNA clone HEMBI1001966 5', mRNA

DEFINITION sequence.
 ACCESSION AU121047
 VERSION AU121047.1 GI:10936282

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 844)
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,D., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

TITLE HRI human cDNA project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3951
 Fax: 81-438-52-3952
 Email: genomics@hri.co.jp

FEATURES
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 Location/Qualifiers

source

1. .844
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEMBI1001966"
 /clone_1ib="HEMBI1"

/tissue_type="whole embryo, mainly body"
 /dev_stage="embryo, 10 weeks"
 /note="Vector: pME185FL3"

BASE COUNT 252 a 179 c 205 g 204 t 4 others

Query Match 100.0%; Score 500; DB 9; Length 844;
 Best Local Similarity 100.0%; Pred. No. 4.8e-117;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAGCTGAGCCGCTGCGCCGCAACCGGGCTAGACCCGCGCCATCATGCTGCTTCCCAA 60
 DB 106 GGGAGCTGAGCCGCTGCGCCGCAACCGGGCTAGACCCGCGCCATCATGCTGCTTCCCAA 165
 QY 61 GGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 DB 166 GGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 225
 QY 121 GTCAAGTTGAACCTGTGGAGAGATGGTTTGAAGCTTTTGTAAAGAGAGAAACAGAAATG 180
 DB 226 GTCAAGTTGAACCTGTGGAGAGATGGTTTGAAGCTTTTGTAAAGAGAGAAACAGAAATG 285
 QY 181 GTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAAGATTAAGAAAGATTAAGAAAGATG 240
 DB 286 GTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAAGATTAAGAAAGATTAAGAAAGATG 345
 QY 241 AAGAAATTCAGTTGGAAGATTTCCCATGCTGAAACACTTGAATCCCAAGACTGGAGA 300
 DB 346 AAGAAATTCAGTTGGAAGATTTCCCATGCTGAAACACTTGAATCCCAAGACTGGAGA 405
 QY 301 ACCAAGATCATATGCAAGTTTGGAGCTTGGCATGTGAGATCAAGGCTACACAGAGAC 360
 DB 406 ACCAAGATCATATGCAAGTTTGGAGCTTGGCATGTGAGATCAAGGCTACACAGAGAC 465
 QY 361 AGATCAAGAGCTCATTAAGCAATGTTTAAACATCAACCCAGAACGAGAAAGCAG 420
 DB 466 AGATCAAGAGCTCATTAAGCAATGTTTAAACATCAACCCAGAACGAGAAAGCAG 525
 QY 421 CTGGTGACCAATTAAGAAAGAGATTAAGTACTTCACTTGCATTAACATTAAGCTTATG 480
 DB 526 CTGGTGACCAATTAAGAAAGAGATTAAGTACTTCACTTGCATTAACATTAAGCTTATG 585
 QY 481 AATGTTATCTGATCCAGTG 500
 DB 586 AATGTTATCTGATCCAGTG 605

RESULT 3

BC012376 920 bp mRNA linear HMC 20-AUG-2001
 LOCUS BC012376 Homo sapiens, clone IMAGE:3860504, mRNA.

DEFINITION BC012376
 ACCESSION BC012376
 VERSION BC012376.1 GI:15214504

KEYWORDS HMC.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 920)
 Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
 Strausberg,R.
 Direct Submission
 Submitted (15-AUG-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalob@bcm.tmc.edu.
 Villalob, D.K., Luna, R.A., Hale, S.M., Buljck, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>
 Series: IRK Plate: 21 Row: e Column: 16
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Genomescan gene prediction, similarity but not identity
 to protein

This clone has the following problem: incomplete processing.

FEATURES

source
 Location/Qualifiers
 1. 920
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3860504"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_lib="NIH_MGC_66"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 BASE COUNT 304 a 180 c 222 g 214 t
 ORIGIN

Query Match 100.0%; Score 500; DB 11; Length 920;
 Best Local Similarity 100.0%; Pred. No. 5e-117;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GGGACGTGAGCCGCTGCGCCACCGGGCTAGACCGCGCCATCATGCTCTTGTGCCAA 60
 DB GGGACGTGAGCCGCTGCGCCACCGGGCTAGACCGCGCCATCATGCTCTTGTGCCAA 148
 89 GGGACGTGAGCCGCTGCGCCACCGGGCTAGACCGCGCCATCATGCTCTTGTGCCAA 148
 61 GCGCGCGGAGCGCGGGGCGACCGCCATCACCAGCTCTGACCTCTGCTTACACTCT 120
 DB GCGCGCGGAGCGCGGGGCGACCGCCATCACCAGCTCTGACCTCTGCTTACACTCT 208
 149 GCGCGCGGAGCGCGGGGCGACCGCCATCACCAGCTCTGACCTCTGCTTACACTCT 208
 121 GTCAATTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGAAGGAGAAACAGAAATG 180
 DB GTCAATTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGAAGGAGAAACAGAAATG 268
 209 GTCAATTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGAAGGAGAAACAGAAATG 268
 181 CTCTGCTCTTTTTCAGAACTGAGAGATPAGAAGAGTTATCCAGAGATCAGAAATG 240
 DB CTCTGCTCTTTTTCAGAACTGAGAGATPAGAAGAGTTATCCAGAGATCAGAAATG 328
 269 CTCTGCTCTTTTTCAGAACTGAGAGATPAGAAGAGTTATCCAGAGATCAGAAATG 328
 241 AAGAAATTTGAGTGGGAAGCTTTCCATGCTGAAAACCTTATCCCAAGACTGGAAGA 300
 DB AAGAAATTTGAGTGGGAAGCTTTCCATGCTGAAAACCTTATCCCAAGACTGGAAGA 388
 329 AAGAAATTTGAGTGGGAAGCTTTCCATGCTGAAAACCTTATCCCAAGACTGGAAGA 388
 301 ACCAAGATCATTTAGAGTCTTGGACTTGGCCATGTGATPACAGAGCTTACACAGAGAC 360
 DB ACCAAGATCATTTAGAGTCTTGGACTTGGCCATGTGATPACAGAGCTTACACAGAGAC 448
 389 ACCAAGATCATTTAGAGTCTTGGACTTGGCCATGTGATPACAGAGCTTACACAGAGAC 448
 361 AGATTCAGAGAGCTCATTAAGCAATGTTTAAACATCACCCAGCAACAGGAAGAG 420
 DB AGATTCAGAGAGCTCATTAAGCAATGTTTAAACATCACCCAGCAACAGGAAGAG 508
 449 AGATTCAGAGAGCTCATTAAGCAATGTTTAAACATCACCCAGCAACAGGAAGAG 508
 421 CTGTGTGACCAATTAAGAGAGATTAATGACTACTCTTGCATTAAGAGCTTATG 480
 DB CTGTGTGACCAATTAAGAGAGATTAATGACTACTCTTGCATTAAGAGCTTATG 568
 509 CTGTGTGACCAATTAAGAGAGATTAATGACTACTCTTGCATTAAGAGCTTATG 568
 481 AAATGTATCTGATCCAGTG 500
 DB AAATGTATCTGATCCAGTG 588

RESULT 4
 BG720251 666 bp mRNA linear EST 08-MAY-2001
 LOCUS 602692311f1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4824555 5',
 DEFINITION mRNA sequence.
 ACCESSION BG720251
 VERSION BG720251.1 GI:13999438
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 666)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiyaki
 Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
 Plate: LLM10736 Row: e Column: 04
 High quality sequence stop: 661.
 Location/Qualifiers
 1. 666
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4824555"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptPR (modified
 pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtagag
); Oligo-dT primed using primer 5'-TTTTTATTTTATTTT-3',
 size-selected for average insert size 2.2 kb and
 normalized to 10^5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

FEATURES

source

BASE COUNT 198 a 155 c 173 g 140 t
 ORIGIN

Query Match 97.8%; Score 489; DB 10; Length 666;
 Best Local Similarity 99.8%; Pred. No. 2.8e-114;
 Matches 500; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 GGGACGTGAGCCGCTGCGCCACCGGGCTAGACCGCGCCATCATGCTCTTGTGCCAA 60
 DB GGGACGTGAGCCGCTGCGCCACCGGGCTAGACCGCGCCATCATGCTCTTGTGCCAA 153
 94 GGGACGTGAGCCGCTGCGCCACCGGGCTAGACCGCGCCATCATGCTCTTGTGCCAA 153
 61 GCGCGCGGAGCGCGGGGCGACCGCCATCACCAGCTCTGACCTCTGCTTACACTCT 120
 DB GCGCGCGGAGCGCGGGGCGACCGCCATCACCAGCTCTGACCTCTGCTTACACTCT 213
 154 GCGCGCGGAGCGCGGGGCGACCGCCATCACCAGCTCTGACCTCTGCTTACACTCT 213
 121 GTCAATTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGAAGGAGAAACAGAAATG 180
 DB GTCAATTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGAAGGAGAAACAGAAATG 273
 214 GTCAATTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGAAGGAGAAACAGAAATG 273
 181 CTCTGCTCTTTTTCAGAACTGAGAGATPAGAAGAGTTATCCAGAGATCAGAAATG 240
 DB CTCTGCTCTTTTTCAGAACTGAGAGATPAGAAGAGTTATCCAGAGATCAGAAATG 333
 274 CTCTGCTCTTTTTCAGAACTGAGAGATPAGAAGAGTTATCCAGAGATCAGAAATG 333
 241 AAGAAATTTGAGTGGGAAGCTTTCCATGCTGAAAACCTTATCCCAAGACTGGAAGA 300
 DB AAGAAATTTGAGTGGGAAGCTTTCCATGCTGAAAACCTTATCCCAAGACTGGAAGA 393
 334 AAGAAATTTGAGTGGGAAGCTTTCCATGCTGAAAACCTTATCCCAAGACTGGAAGA 393
 301 ACCAAGATCATTTAGAGTCTTGGACTTGGCCATGTGATPACAGAGCTTACACAGAGAC 360

QY 1 GGGAGCTGAGCCGCTGCGCCGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 60
 DB 72 GGGAGCTGAGCCGCTGCGCCGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 131
 QY 61 GCGCGCGGACGCGCGGCGGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 120
 DB 132 GCGCGCGGACGCGCGGCGGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 191
 QY 121 GTCAGTGAACCTGTGGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 180
 DB 192 GTCAGTGAACCTGTGGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 250
 QY 181 CTCTGCTCTCTTTCAGGAATGAGATAGAGAAAGATTATCCGAGGAATCAGAAATG 240
 DB 251 CTCTGCTCTCTTTCAGGAATGAGATAGAGAAAGATTATCCGAGGAATCAGAAATG 310
 QY 241 AAGAAATGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 300
 DB 311 AAGAAATGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 370
 QY 301 ACCAAGATCATTTAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 360
 DB 371 ACCAAGATCATTTAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 430
 QY 361 AGATCAAGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 420
 DB 431 AGATCAAGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 490
 QY 421 CTGTGAGACCAATTAAGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 480
 DB 491 CTGTGAGACCAATTAAGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 550
 QY 481 AATGTTATCTGATCCAGTG 500
 DB 551 AATGTTATCTGATCCAGTG 570

RESULT 7 937 bp mRNA linear EST 12-MAR-2001
 LOCUS BC396661
 DEFINITION 602459730F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4582038 5',
 mRNA sequence.
 ACCESSION BC396661
 VERSION BC396661.1 GI:13290109
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 937)
 AUTHORS NIH-MGC, <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LHC41303 row: 1 column: 07
 High quality sequence stop: 758.
 location/Qualifiers
 1. 937
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4582038"
 /clone_lib="NIH_MGC_16"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:

BASE COUNT 280 a 192 c 244 g 221 t
 ORIGIN
 Query Match 97.6%; Score 488; DB 10; Length 937;
 Best Local Similarity 99.8%; Pred. No. 5,8e-114;
 Matches 499; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 GGGAGCTGAGCCGCTGCGCCGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 60
 DB 86 GGGAGCTGAGCCGCTGCGCCGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 145
 QY 61 GCGCGCGGACGCGCGGCGGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 120
 DB 146 GCGCGCGGACGCGCGGCGGACGCGGCTAGACCCGCGCCATCATGCTTTCGCCAA 205
 QY 121 GTCAGTGAACCTGTGGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 180
 DB 206 GTCAGTGAACCTGTGGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 265
 QY 181 CTCTGCTCTCTTTCAGGAATGAGATAGAGAAAGATTATCCGAGGAATCAGAAATG 240
 DB 266 CTCTGCTCTCTTTCAGGAATGAGATAGAGAAAGATTATCCGAGGAATCAGAAATG 325
 QY 241 AAGAAATGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 300
 DB 326 AAGAAATGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 385
 QY 301 ACCAAGATCATTTAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 360
 DB 386 ACCAAGATCATTTAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 444
 QY 361 AGATCAAGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 420
 DB 445 AGATCAAGAGAGATGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 504
 QY 421 CTGTGAGACCAATTAAGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 480
 DB 505 CTGTGAGACCAATTAAGAGAGATGCTTGAAGCTTTGTTAAGAGAGAGAGAAATG 564
 QY 481 AATGTTATCTGATCCAGTG 500
 DB 565 AATGTTATCTGATCCAGTG 584

RESULT 8
 LOCUS AL517760 955 bp mRNA linear EST 13-FEB-2001
 DEFINITION AL517760 LRL_NFL011.NBC1 Homo sapiens cDNA clone CS00A004YR24 5
 prime, mRNA sequence.
 ACCESSION AL517760
 VERSION AL517760.1 GI:12781253
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 955)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 101 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 location/Qualifiers
 1. 955


```

/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="CS00AD004E24"
/clone.lib="LRI_NF011_NBC1"
/sex="male"
/tissue="neuroblastoma cells"
/lab.host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax: (1) 301 610 8371 Email:
f.liang@lifestech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT      328 a      174 c      222 g      221 t      10 others
ORIGIN
Query Match      97.3%; Score 486.6; DB 9; Length 955;
Best Local Similarity 98.6%; Pred. No. 1,3e-113;
Matches 494; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGGACGTGAGCCGCTGCGCCACCGGCTAGACCGGCGCCATCATGCTTCTGCCAA 60
    |||||
DB 49 GGGACGTGAGCCGCTGCGCCACCGGCTAGACCGGCGCCATCATGCTTCTGCCAA 108
    |||||

QY 61 GCGCGCGCGG-ACGGCGGGGGCGACCGCCATCCAGCCGCTGCTGCTTACACTC 119
    |||||
DB 109 GCGCGCGGGAGCGCGGGGGCGACCGCCATCCAGCCGCTGCTGCTTACACTC 168
    |||||

QY 120 TGTCAAGTTGAACTGTGGAGAAAGGTTTGAAGCTTTTGTAGAGGAGAAACAAT 179
    |||||
DB 169 TGTCAAGTTGAACTGTGGAGAAAGGTTTGAAGCTTTTGTAGAGGAGAAACAAT 228
    |||||

QY 180 GCTTGTGCTCTTTTTCAGAGACTGAGATTAAGAAAGTATCCAGAGAAATCAAAAT 239
    |||||
DB 229 GCTTGTGCTCTTTTTCAGAGACTGAGATTAAGAAAGTATCCAGAGAAATCAAAAT 288
    |||||

QY 240 GAAGAATTGCACTTGGAGAGATTTCCTTCACTGTAAGAACTTGAATCCCAAGCTGAG 299
    |||||
DB 289 GAAGAATTGCACTTGGAGAGATTTCCTTCACTGTAAGAACTTGAATCCCAAGCTGAG 348
    |||||

QY 300 AACCAAGATCATTAATGAGATTTCTTGACTTGGCCATGTGATTAACAAGCTACAGAGA 359
    |||||
DB 349 AACCAAGATCATTAATGAGATTTCTTGACTTGGCCATGTGATTAACAAGCTACAGAGA 408
    |||||

QY 360 CAGATCAAAAGCAGCTCATAAAGCAATGTTTTTAAACATCAACCCAGCAAAAGCAAGA 419
    |||||
DB 409 CAGATCAAAAGCAGCTCATAAAGCAATGTTTTTAAACATCAACCCAGCAAAAGCAAGA 468
    |||||

QY 420 GCTGTGGAACCAATAAAGAGAGATTAATGACTTCACTTCACTTCACTTCACTTAT 479
    |||||
DB 469 GCTGTGGAACCAATAAAGAGAGATTAATGACTTCACTTCACTTCACTTCACTTAT 528
    |||||

QY 480 GAAATGTATCTGATCCAGTG 500
    |||||
DB 529 GAAATGTATCTGATCCAGTG 549
    |||||

RESULT 9
LOCUS      BG718443      829 bp      mRNA      linear      EST 08-MAY-2001
DEFINITION 602696473P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828339 5',
ACCESSION  BG718443
VERSION     BG718443.1  GI:13997630
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE 1 (bases 1 to 829)
AUTHORS   NIH-NCBI http://imgc.ncbi.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs@emall.nih.gov
          Tissue Procurement: Miklos Palkovics, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NIGRI), Shitaki
          Yoshiyuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LLM410746 row: D column: 20
          High quality sequence stop: 822.
          Location/Qualifiers
            1..829
            /organism="Homo sapiens"
            /db.xref="taxon:9606"
            /clone="IMAGE:4828339"
            /clone.lib="NIH_MGC_97"
            /lab.host="DH10B"
            /note="Organ: testis; Vector: pBluescript (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTTTNNNNNNNNNN-3',
            size-selected for average insert size 2.2 kb and
            normalized to 10^5. This is a primary library enriched
            for full-length clones and constructed using the
            cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NIMH/NHRI, National
            Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT      263 a      165 c      201 g      199 t      1 others
ORIGIN
Query Match      97.1%; Score 485.4; DB 10; Length 829;
Best Local Similarity 99.4%; Pred. No. 2,3e-113;
Matches 497; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 GGGACGTGAGCCGCTGCGCCACCGGCTAGACCGGCGCCATCATGCTTCTGCCAA 60
    |||||
DB 58 GGGACGTGAGCCGCTGCGCCACCGGCTAGACCGGCGCCATCATGCTTCTGCCAA 117
    |||||

QY 61 GCGCGCGGGAGCGCGGGGGCGACCGCCATCCAGCCGCTGCTGCTTACACTCT 120
    |||||
DB 118 GCGCGCGGGAGCGCGGGGGCGACCGCCATCCAGCCGCTGCTGCTTACACTCT 176
    |||||

QY 121 GTCAAAGTTGAACCTGTGGAGAAAGTGTGAAGCTTTTGTAGAGAGAAACAATG 180
    |||||
DB 177 GTCAAAGTTGAACCTGTGGAGAAAGTGTGAAGCTTTTGTAGAGAGAAACAATG 236
    |||||

QY 181 CTTCTGCTCTTTTTCAGAACTGAGATTAAGAAAGATTCAGAGAAATG 240
    |||||
DB 237 CTTCTGCTCTTTTTCAGAACTGAGATTAAGAAAGATTCAGAGAAATG 296
    |||||

QY 241 AAGAAATGCAAGTTGGAAGAGTTTCCATGCTGAAGAACATCTGATCCCAAGACTGGAAGA 300
    |||||
DB 297 AAGAAATGCAAGTTGGAAGAGTTTCCATGCTGAAGAACATCTGATCCCAAGACTGGAAGA 356
    |||||

QY 301 ACCAAGATCATTAATGAGATTTCTTGACTTGGCCATGTGATTAACAAGCTACAGAGA 360
    |||||
DB 357 ACCAAGATCATTAATGAGATTTCTTGACTTGGCCATGTGATTAACAAGCTACAGAGA 416
    |||||

QY 361 AGATCAAGAGAGCTCATAAAGAGTGTAAAGATCAACCCAGCAAAAGCAAGAGAG 420
    |||||
DB 417 AGATCAAGAGAGCTCATAAAGAGTGTAAAGATCAACCCAGCAAAAGCAAGAGAG 476
    |||||

QY 421 CTGGTGAAACCAATAAAGAGAGATTAATGACTTCACTTCACTTCACTTAT 480
    |||||
DB 477 CTGGTGAAACCAATAAAGAGAGATTAATGACTTCACTTCACTTCACTTAT 536
    |||||

QY 481 AATGTATCTGATCCAGTG 500
    |||||

```


QY 181 CTTTCGCCCTTTTTCAGAACTGAGATTAAGAAAGAGTTATCCGAGGAATCAGAAAGATG 240
 |||||||
 Db 180 CTTTCGCCCTTTTTCAGAACTGAGATTAAGAAAGAGTTATCCGAGGAATCAGAAAGATG 239
 |||||||
 QY 241 AAGAAATGAGTTGGAGAGTTTCCATGCTGAAAACCTTGCATCCCAAGCAAGCTGAAAGA 300
 |||||||
 Db 240 AAGAA-TGCAAGTTGGAGAGTTTCCATGCTGAAAACCTTGCATCCCAAGCAAGCTGAAAGA 298
 |||||||
 QY 301 ACCAAGATCATTAATGAGTCTTTGGAGCTTGGCCATGATGATACAAAGGCTACACAGAGAC 360
 |||||||
 Db 299 ACCAAGATCATTAATGAGTCTTTGGAGCTTGGCCATGATGATACAAAGGCTACACAGAGAC 358
 |||||||
 QY 361 AAGTCAAGCAGCTGATTAAGCAATGTTTAAACATCACCCAGACAAAGCAAGAGAGAG 420
 |||||||
 Db 359 AAGTCAAGCAGCTGATTAAGCAATGTTTAAACATCACCCAGACAAAGCAAGAGAGAG 418
 |||||||
 QY 421 CTTGTGAACCATTAAGAGAGATTAATGACTACTTCACTTGCATTAAGCTTAAGCTTAATG 480
 |||||||
 Db 419 CTTGTGAACCATTAAGAGAGATTAATGACTACTTCACTTGCATTAAGCTTAAGCTTAATG 478
 |||||||
 QY 481 AATGTATATGATCCAGTG 500
 |||||||
 Db 479 AATGTATATGATCCAGTG 498
 |||||||

RESULT 12
 AT364977/c 735 bp mRNA linear EST 16-FEB-1999
 LOCUS AT364977

DEFINITION q241h10.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2029507 3' similar to TR:099543 099543 M-PHASE PHOSPHOPROTEIN 11 ; mRNA sequence.

ACCESSION AT364977
 VERSION AT364977.1 GI:4124666
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 735)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 TITLE Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: gcgapsb@mail.nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrrp/image/image.html
 Insert length: 572 Std Error: 0.00
 Seq primer: 400P from Gibco
 High quality sequence stop: 371.
 Location/Qualifiers

FEATURES

source

1. 735
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2029507"
 /clone_lib="NCI-CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: Kidney; Vector: pT773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M.

BASE COUNT 194 a 162 c 144 g 233 t 2 others
 ORIGIN
 Query Match 94.6%; Score 472.8; DB 9; Length 735;
 Best local Similarity 98.0%; Pred No. 4e-110;
 Matches 488; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 3 GACGTGAGCGGCTGCGCCACCGGGGCTAGACCGGGCCCATCATCTCTCTTGGCAAGC 62
 |||||||
 Db 735 GACGTGAGCGGCTGCGCCACCGGGGCTAGACCGGGCCCATCATCTCTTGGCAAGC 676
 |||||||
 QY 63 GCGCGAGCGCGCGGGGCGACCCGCATCACCCAGCTTGCCTTACACTCTGT 122
 |||||||
 Db 675 GCGCGAGCGCGGGGCGACCCGCATCACCCAGCTTGCCTTACACTCTGT 616
 |||||||
 QY 123 CAAGTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGTAAAGAGAAACAGAAATGCT 182
 |||||||
 Db 615 CAAGTTGAACCTGTGGGAAGATGTTTGAAGCTTTTGTAAAGAGAAACAGAAATGCT 556
 |||||||
 QY 183 TCTGCCCTTTTCAGGAGACTGAGATTAAGAAAGATTATCCGAGAAATCAGAAATGAA 242
 |||||||
 Db 555 TCTGCCCTTTTCAGGAGACTGAGATTAAGAAAGATTATCCGAGAAATCAGAAATGAA 496
 |||||||
 QY 243 GAATTCAGTTGGAAGAGTTTCCCATGCTGAAGAAACACTTGAATCCCAAGAGCTGGAAGAC 302
 |||||||
 Db 495 GAATTCAGTTGGAAGAGTTTCCCATGCTGAAGAAACACTTGAATCCCAAGAGCTGGAAGAC 437
 |||||||
 QY 303 CAAGATCATTAATGAGTCTTTGGAGCTTGGCCATGATGATACAAAGGCTACACAGAGACG 362
 |||||||
 Db 436 CAAGATCATTAATGAGTCTTTGGAGCTTGGCCATGATGATACAAAGGCTACACAGAGACG 377
 |||||||
 QY 363 ATCAAGCAGCTCATTAAGCAATGTTTAAACATCACCCAGACAAAGCAAGAGACGCT 422
 |||||||
 Db 376 ATCAAGCAGCTCATTAAGCAATGTTTAAACATCACCCAGACAAAGCAAGAGACGCT 317
 |||||||
 QY 423 GGTGAACCAATTAAGAGAGATTAATGACTTCACTTGCATTAAGCTTAATGAA 482
 |||||||
 Db 316 GGTGAACCAATTAAGAGAGATTAATGACTTCACTTGCATTAAGCTTAATGAA 257
 |||||||
 QY 483 ATGTATATGATCCAGTG 500
 |||||||
 Db 256 ATGTATATGATCCAGTG 239
 |||||||

RESULT 13
 AM161245 579 bp mRNA linear EST 09-NOV-1999
 LOCUS AM161245

DEFINITION au76d01.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone IMAGE:2781601 5' similar to TR:060415 060415 WUGSC:H_RG276003.1A PROTEIN ; mRNA sequence.

ACCESSION AM161245
 VERSION AM161245.1 GI:6300278
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 579)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Merri, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theisling, B., White, X., Wyllie, T., Waterston, R. and Wilson, R.
 WASHU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Other ESTs: au76d01.x1
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.jhu.gov) for further information.
Seq primer: -40RP from Gibco
High quality sequence stop: 417.

FEATURES
Source

1. 579
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2781601"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site 1: SstI; Site 2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAAGAGAGCTCAGAGATCCTTAATTAATTAATTCCTCCCTCCCTCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGCTCAGAGATCCTTAATTAATTAATTCCTCCCTCCCTCC-3'
size selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."

BASE COUNT 162 a 139 c 157 g 121 t
ORIGIN

Query Match 94.4%; Score 472; DB 9; Length 579;

Best Local Similarity 97.8%; Pred. No. 5.8e-110; Mismatches 10; Indels 1; Gaps 1;

Matches 489; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

1 GGGAGCTGAGCGCGCTGCGCCACCGGGCTAGACCGGGCCATCATGCTCTTGGCCAA 60
|||||
81 GGGAGCTGAGCGCGCTGCGCCACCGGGCTAGACCGGGCCATCATGCTCTTGGCCAA 139
|||||
61 GGGCGCGGAGCGCGGGGCGCCCATCACCGCGCTGACCTTGGCTCTACACTCT 120
|||||
140 GGGCGCGGAGCGCGGGGCGCCCATCACCGCGCTGACCTTGGCTCTACACTCT 199
|||||
121 GTCAGTTGAACCTGTGGAGAGTGTGGAAGCTTTGTTAAGAGAGAAACAGAAATG 180
|||||
200 GTCAGTTGAACCTGTGGAGAGTGTGGAAGCTTTGTTAAGAGAGAAACAGAAATG 259
|||||
181 CTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAGAGTTATCCAGAGAAATG 240
|||||
260 CTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAGAGTTATCCAGAGAAATG 319
|||||
241 AAGATTTGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 300
|||||
320 AAGATTTGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 379
|||||
301 ACCAAGATCATTTATGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 360
|||||
380 ACCAAGATCATTTATGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 439
|||||
361 AGATCAAGAGAGCTCATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAGCAAG 420
|||||
440 AGATCAAGAGAGCTCATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAGCAAG 499
|||||
421 CTGTGTAACCAATTAAGCAAGAGATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAG 480
|||||
500 CTGTGTAACCAATTAAGCAAGAGATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAG 559
|||||
481 AAATGTTATCTGATCAAGT 500
|||||
560 GAAATTTATCTGATCAAGT 579
|||||

RESULT 14
A0128589

LOCUS A0128589 721 bp mRNA linear EST 24-OCT-2000
DEFINITION A0128589 NT2RP2 Homo sapiens cDNA clone NT2RP2003763 5', mRNA
sequence.
ACCESSION A0128589
VERSION A0128589.1 GI:10988943
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Isogai,T.
HRI human cDNA project
Unpublished (2000)
CONTACT: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3851
Fax: 81-438-52-3852
Email: genomics@hri.co.jp

TITLE
JOURNAL
COMMENT
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
Source

Location/Qualifiers

1. 721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2003763"
/clone_lib="NT2RP2"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 209 a 171 c 170 g 164 t
ORIGIN

Query Match 92.1%; Score 460.6; DB 9; Length 721;

Best Local Similarity 96.8%; Pred. No. 5.1e-107; Mismatches 15; Indels 1; Gaps 1;

Matches 480; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

1 GGGAGCTGAGCGCGCTGCGCCACCGGGCTAGACCGGGCCATCATGCTCTTGGCCAA 60
|||||
89 GGGAGCTGAGCGCGCTGCGCCACCGGGCTAGACCGGGCCATCATGCTCTTGGCCAA 148
|||||
61 GCGCGCGGAGCGCGGGGCGCCCATCACCGCGCTGACCTTGGCTCTACACTCT 120
|||||
149 GCGCGCGGAGCGCGGGGCGCCCATCACCGCGCTGACCTTGGCTCTACACTCT 208
|||||
121 GTCAGTTGAACCTGTGGAGAGTGTGGAAGCTTTGTTAAGAGAGAAACAGAAATG 180
|||||
209 GTCAGTTGAACCTGTGGAGAGTGTGGAAGCTTTGTTAAGAGAGAAACAGAAATG 268
|||||
181 CTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAGAGTTATCCGAGAGATCAGAAATG 240
|||||
269 CTTCTGCTCTTTTTCAGAACTGGAGATTAAGAAGAGTTATCCGAGAGATCAGAAATG 328
|||||
241 AAGATTTGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 300
|||||
329 AAGATTTGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 388
|||||
301 ACCAAGATCATTTATGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 360
|||||
389 ACCAAGATCATTTATGCAAGTTGGAAGATTTCCCATGCTGAAAACACTGTATCCCAAGACTGGAGA 448
|||||
361 AGATCAAGAGAGCTCATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAGCAAG 419
|||||
449 AGATCAAGAGAGCTCATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAGCAAG 508
|||||
420 GCTGTGAACCAATTAAGCAAGAGATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAG 479
|||||

Db 509 GCTGTCGACCAATTAAGAAGAGATTAATGACTTCACTTGCATTAAGTATGATCC 568
 QY 480 GAAATGTTATCTGATC 495
 Db 569 CCAAGTCATTTCTTATC 584

RESULT 15
 BI562033 656 bp mRNA linear EST 05-SEP-2001
 LOCUS 603255091F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5297303 5',
 DEFINITION mRNA sequence.

ACCESSION BI562033
 VERSION BI562033.1 GI:15449347
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 656)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue procurement: Miklos Palxovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLAM1752 row: f column: 24
 high quality sequence stop: 654.
 Location/Qualifiers

FEATURES
 source
 1. 656
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5297303"
 /clone_1ib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 184 a 169 c 173 g 130 t
 ORIGIN

Query Match 87.6%; Score 438.2; DB 10; Length 656;
 Best Local Similarity 96.8%; Pred. No. 2.6e-101;
 Matches 479; Conservative 0; Mismatches 13; Indels 3; Gaps 3;

QY 1 GGGACGTAGCCCGCTGGCCCGCCAGCCGCTAGACCCGCCCATCATGCTGCTTGCCTAA 60
 Db 152 GGGACGTAGCCCGCTGGCCCGCCAGCCGCTAGACCCGCCCATCATGCTGCTTGCCTAA 211
 QY 61 GCGCCGGGAGCGCGGGGACCGCCATCACACGCTGACCTTGCCCTTACACTCT 120
 Db 212 GCGCCGGGAGCGGCGGGGACCGCCATCACACGCTGACCTTGCCCTTACACTCT 270
 QY 121 GTCAAGTTGAACCTGTTGGAGAGATGTTTGAAGCTTTTGAAGAGAGAGAAACAGAAATG 180
 Db 271 GTCAAGTTGAACCTGTTGGAGAGATGTTTGAAGCTTTTGAAGAGAGAGAAACAGAAATG 329
 QY 181 CTTCCTGCTCTTTTACAGAACTGGAGATTAAGAAAGATTATCCGAGGAATCAGAGATG 240

Db 330 CTTCCTGCTCTTTTACAGAACTGGAGATTAAGAAAGATTATCCGAGGAATCAGAGATG 389
 QY 241 AAGAAATGAGTTGGAGAGATTTCCATGCGTGAAGAACCTTGATCCCAAGACTGGAAGA 300
 Db 390 AAGAA-TGCAGTTGGAAAGATTTCCATGCGTGAAGAACCTTGATCCCAAGACTGGAAGA 448
 QY 301 ACCAAGATCATTTATGCACTTCTTGACCTTGGCCATGTGATACAGGCTTACAGAGAC 360
 Db 449 ACCAAGATCATTTATGCACTTCTTGACCTTGGCCATGTGATACAGGCTTACAGAGAC 508
 QY 361 AGATCAAGACAGCTCATTAAGCAATGTTTAAACATCACCCAGACAAACGGAAGACAG 420
 Db 509 AGATCAAGACAGCTCATTAAGCAATGTTTAAACATCACCCAGACAAACGGAAGACAG 568
 QY 421 CTGCTGAACCAATTAAGAGAGATTAATGACTTCACTTGCATTAACCTTATG 480
 Db 569 CTGCTGAACCAATTAAGAGAGATTAATGACTTCACTTGCATTAACCTTATG 628
 QY 481 AAATGTTATCTGATC 495
 Db 629 CAGGTGATTTCTTATC 643

Search completed: October 10, 2002, 20:27:42
 Job time : 702.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 84.933 Seconds
(without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101a-8_COPY_1_500
Perfect score: 500
Sequence: 1 gggacgtgagcgcgtgcgc.....aaatgtatctgcatcagtg 500

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001C.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	876	21	AAC98111
2	500	100.0	1845	24	ABK09776
3	500	100.0	1860	22	ABD11115
4	80.2	16.0	2225	23	ABL16647
5	60.4	12.1	1876	22	AAH29904
6	51.8	10.4	1802	22	AAH29775
7	51.8	10.4	3083	18	AAH90969
8	51.8	10.4	3083	20	AA211218
9	42.4	8.5	4791	23	ABL16646

C	10	42.4	8.5	5071	23	ABL16648	Drosophila melanog
	11	42.4	8.2	580073	18	AAH58840	Mycoplasma genital
	12	41.2	8.2	1039	20	AAH61812	B. burgdorferi ant
	13	41.2	8.2	1125	20	AAH61811	B. burgdorferi ant
C	14	41.2	8.2	35515	20	AAH20252	Borrelia burgdorfe
	15	40.8	8.2	3399	17	AAH05868	Chicken leucocytos
	16	40.2	8.0	2700	22	AAH17193	ERP48 insert DNA i
	17	40.2	8.0	65140	22	AAH17184	Streptomyces nous
	18	40.2	8.0	125401	22	AAH17186	Streptomyces nous
	19	38.4	7.7	843	23	AAH90709	DNA encoding novel
	20	36.8	7.4	1686	16	AAH97587	DNA encoding novel
C	21	36	7.2	1381	22	AAH32674	Human genomic DNA
C	22	36	7.2	1384	22	AAH32671	Human genomic DNA
C	23	36	7.2	1386	22	AAH32675	Human genomic DNA
C	24	35.8	7.2	330	22	AAH64823	Human immune/haema
C	25	35.8	7.2	330	22	AAH64824	Human immune/haema
C	26	35.8	7.2	330	22	AAH64825	Human immune/haema
C	27	35.8	7.2	1305	22	AAH06649	Drosophila melanog
C	28	35.8	7.2	3935	23	ABL06648	Drosophila melanog
C	29	35.4	7.1	1510	17	AAH29628	DNA mismatch repair
C	30	35.4	7.1	1510	17	AAH29628	Human strimal anti
	31	35.4	7.1	58857	21	AAH58471	Nucleotide sequenc
C	32	35.2	7.0	377	16	AAH90828	Wild type hMLH1 ge
	33	35.2	7.0	3579	21	AAH70099	Plasmodium falcipla
	34	35.2	7.0	38734	20	AAH232020	Human MERT1 relate
	35	35.2	7.0	38734	22	AAH90077	Al021529 CDNA clon
	36	35	7.0	450	11	AAH03633	Mycoplasma hyopneu
	37	35	7.0	1125	21	AAH56820	Mycoplasma hyopneu
	38	35	7.0	4056	21	AAH70225	Plasmodium falcipla
	39	34.8	7.0	1293	20	AAH30135	Streptococcus pneu
	40	34.8	6.9	910715	20	AAH20248	Novel human diagno
	41	34.6	6.9	393	22	AAH39178	Streptococcus pneu
	42	34.6	6.9	1924	19	AAH27403	Plasmodium falcipla
	43	34.6	6.9	2223	16	AAH080908	S. pneumoniae defi
C	44	34.4	6.9	1372	19	AAH96236	Human CDNA encodin
	45	34.4	6.9	1565	22	AAH26008	

ALIGNMENTS

RESULT 1	
AAC98111	
ID	AAC98111 standard; CDNA: 876 BP.
AC	AAC98111;
XX	
DT	09-MAR-2001 (first entry)
XX	
DE	Human colon cancer antigen nucleotide sequence SEQ ID NO:121.
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW	identification; cytostatic; cardioactive; neuroprotective; vulnary;
KW	immunomodulatory; muscular; gynaecological; gastrointestinal;
KW	neurotropic; antiinfective; antibacterial; gene therapy; wound;
KW	reproductive disorder; immune system disorder; muscular disorder;
KW	infectious disease; cardiovascular disorder; renal disorder;
OS	Homo sapiens.
XX	
PN	WO20005351-A1.
XX	
PD	21-SEP-2000.
XX	
PF	08-MAR-2000; 2000WO-US05883.
XX	
PR	12-MAR-1999; 99US-0124270.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruben SM;
XX	

DB 181 CTCTGCTCTTTTCAAGAACTGGAGATTAAGAAAGACTTATCCGAGCAATCAGACATG 240
 OY 241 AAGAAATTCAGATTGGAAGAGTTTCCCATGCTGAAAAACATTGATCCCAAGACGTGAAGA 300
 CC |||||||
 CC |||||||
 DB 241 AAGAAATTCAGATTGGAAGAGTTTCCCATGCTGAAAAACATTGATCCCAAGACGTGAAGA 300
 OY 301 ACCAAGATCATATATGACATGCTTGGACCTGGCCATGAGATATACAGGCTACACAGAGAC 360
 DB |||||||
 DB 301 ACCAAGATCATATATGACATGCTTGGACCTGGCCATGAGATATACAGGCTACACAGAGAC 360
 OY 361 AGATCAAAAGCAGCTCATATAAGCAATGGTTTAAACATCCACCACACAAACGAAAGCAG 420
 DB |||||||
 DB 361 AGATCAAAAGCAGCTCATATAAGCAATGGTTTAAACATCCACCACACAAACGAAAGCAG 420
 OY 421 CTGGTGAACCAATTAAGAGAGATATGACTACTTCACTTGCATTAACCTTAAGCTTATG 480
 DB |||||||
 DB 421 CTGGTGAACCAATTAAGAGAGATATGACTACTTCACTTGCATTAACCTTATG 480
 OY 481 AATGTTATCTGATCCAGTG 500
 DB |||||||
 DB 481 AATGTTATCTGATCCAGTG 500

RESULT 3

ID AAD11115 standard; DNA: 1860 BP.

XX AAD11115;

DT 24-SEP-2001 (first entry)

XX Human small cell lung cancer associated gene, MPP11.

KW Human; small cell lung cancer; therapy; hCAP; nucleic acid; NA;
 KW melanoma; cancer; colon; breast; head; neck; transitional cancer;
 KW leiomyosarcoma; helix-loop-helix; HLH; protein-binding factor; MPP11;
 KW synovial sarcoma; cytosolic; ds.

XX Homo sapiens.

PN W020015349-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US02015.

XX 21-JAN-2000; 2000US-0489101.

XX (LUDWIG-) LUDWIG INST CANCER RES.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX (CORR) CORNELL RES FOUND INC.

XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;

XX WPI; 2001-457597/49.

XX Isolated polypeptide, used to treat or prognose a disorder

XX characterized by expression of a hCAP e.g. cancer, is encoded by an

XX isolated nucleic acid comprising an NA Group 3 or 4 molecule -

XX Claim 57; Page 94-95; 152pp; English.

XX The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterised by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterised by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterised by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are small and non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.

CC The present sequence is a DNA encoding human helix-loop-helix (HLH)
 CC protein-binding factor MPP11. This small cell lung cancer associated
 CC gene is designated as NY-SCLC-6.
 XX
 SO Sequence 1860 BP; 738 A; 312 C; 413 G; 397 T; 0 other;

Query Match 100.0%; Score 500; DB 22; Length 1860;
 Best Local Similarity 100.0%; Pred. No. 1,9e-142;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGAGGTGACCGCTGCGCCACCGGGCTAGACCCGGCCATCATGCTGTTGCCAA 60
 DB |||||||
 DB 1 GGGAGGTGACCGCTGCGCCACCGGGCTAGACCCGGCCATCATGCTGTTGCCAA 60
 OY 61 GCGCGCGGAGCGCGCGGGGACCGCATCACCAGCCTGTGACCTGTGACTACTT 120
 DB |||||||
 DB 61 GCGCGCGGAGCGCGCGGGGACCGCATCACCAGCCTGTGACCTGTGACTACTT 120
 OY 121 GTCAGATTGAACCTGTGGGAAGATGTTTGAAGCTTTGTTAAGAGAGAAACAGAAATG 180
 DB |||||||
 DB 121 GTCAGATTGAACCTGTGGGAAGATGTTTGAAGCTTTGTTAAGAGAGAAACAGAAATG 180
 OY 181 CTCTGCTCTTTTCAAGAACTGGAGATTAAGAAAGCTTATCCGAGAAATCAGAAATG 240
 DB |||||||
 DB 181 CTCTGCTCTTTTCAAGAACTGGAGATTAAGAAAGCTTATCCGAGAAATCAGAAATG 240
 OY 241 AAGAAATTCAGATTGGAAGAGTTTCCCATGCTGAAAAACATTGATCCCAAGACGTGAAGA 300
 DB |||||||
 DB 241 AAGAAATTCAGATTGGAAGAGTTTCCCATGCTGAAAAACATTGATCCCAAGACGTGAAGA 300
 OY 301 ACCAAGATCATATATGACATGCTTGGACCTGGCCATGAGATATACAGGCTACACAGAGAC 360
 DB |||||||
 DB 301 ACCAAGATCATATATGACATGCTTGGACCTGGCCATGAGATATACAGGCTACACAGAGAC 360
 OY 361 AGATCAAAAGCAGCTCATATAAGCAATGGTTTAAACATCCACCACACAAACGAAAGCAG 420
 DB |||||||
 DB 361 AGATCAAAAGCAGCTCATATAAGCAATGGTTTAAACATCCACCACACAAACGAAAGCAG 420
 OY 421 CTGGTGAACCAATTAAGAGAGATATGACTACTTCACTTGCATTAACCTTATG 480
 DB |||||||
 DB 421 CTGGTGAACCAATTAAGAGAGATATGACTACTTCACTTGCATTAACCTTATG 480
 OY 481 AATGTTATCTGATCCAGTG 500
 DB |||||||
 DB 481 AATGTTATCTGATCCAGTG 500

RESULT 4

ID ABL16647 standard; DNA: 2225 BP.

XX ABL16647;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1414.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX W0200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PT leading to programmed cell death, useful for treating proliferative disorders, yeast and fungal infections, or for preventing apoptosis in certain diseases -

Claim 1; Fig 1; 218pp; English.

CC The present invention provides the protein and coding sequences of a
CC number of apoptosis associated proteins from the yeast *Saccharomyces*
CC *cerevisiae* and the fungus *Candida albicans*. These can be used to identify
CC treatments for fungal and yeast infections, for proliferative diseases
CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
CC and neurodegeneration. The present sequence is one of the *S. cerevisiae*
CC coding sequences of the invention
CC

SQ Sequence 1802 BP; 593 A; 346 C; 368 G; 495 T; 0 other;

Query Match	10.4%;	Score 51.8;	DB 22;	Length 1802;
Best Local Similarity	55.7%;	Pred. No. 2.3e+05;		
Matches 123;	Conservative 0;	Mismatches 92;	Indels 6;	Gaps 1.

Qy	276	ACACTGATGCCAAGAACTGGAAACACCAAGATCATTAATGACAGTCTTGACATTGGACCTTGGACAT	3350
Db	759	ACACATGATGCTAGAGACTGGAAAACTGCCGATTTTGATGCTGCTATGGGTTTGCTCTAG	8188
Qy	336	GTCGATPACAAAGGTACACACAGACAGTACATCAAAAGCAGCTCATTAAGACCAATGTTTAAAA	3395
Db	819	TTGCGTTTACAGAGCTACAGCAAGTCAATATCATCAAGGCTCACGAAACCAACAGTGTCTCAG	8788
Qy	396	CATCACCCAGACAAACGGAAAGCAGCTGTGGCAACCAATATAAAGAAAGAGATTAAGACTAC	4555
Db	879	TACATCTCCACACAGCAATCTGCTGCTGGTGAGTAAGTTGGACCAAG-----ATGGCTTT	9322
Qy	456	TTGCATTCGATTACTTAAGGCTTATGAAGAAATGTTATCTGATCC	496
Qy	933	TTTCAAGATTATTCAAAGAGCCCTTTGAAACCTTGACGATTC	973

RESULT 7
AAT90969
ID AAT90969 standard; DNA; 3083 BP.

AC	AAT90969;
XX	
DT	22-JAN-1998 (first entry)

KW Membrane-forming peptide; amphiphilic; monovalent metal cation; suture;
 KW internal lining; slow-diffusion drug delivery vehicle; erythropoietin;
 KW tissue-type plasminogen activator; haemoglobin; insulin; artificial skin
 KW separation matrix; dialysis membranes; filter; collagen; cell migration;
 KW Alzheimer's disease; scrapie infection; therapy; proteolytic digestion;
 KW acidic pH resistant; stomach acid; cell adhesion; cell monolayer; zucchini
 KW tissue culture; differentiated cell; stratified cell layer; ds.

Saccharomyces cerevisiae.

FH	Key	Location/Qualifiers
FT	CDS	1292..2593
FT		/*tag= a
FT		/product= zuotin
..		

PI Holmes T, Lockshin C, Rich A, Zhang S, Xu

DR WPI; 1997-479506/44.
DR P-PSDB; AAW30256.

PT Membranes formed by self-assembly of amphiphilic peptide(s) - useful
PT as bio:material(s), separation matrices, drug delivery vehicles,
PT etc.

PS Example 5; Column 29-34; 49pp; English.

This sequence represents the yeast znotin genes. Fragments of the encoded protein can be used as the amphiphilic peptides used in the macroscopic membrane (MM) of the invention. The MM is formed by self-assembly of the amphiphilic peptides (see AA030219-W03033) in an aqueous medium containing monovalent metal cations. The MM are stable in serum, are non-cytotoxic, and are useful in biomaterial applications, such as medical products (e.g. sutures), or internal linings. The MM are useful as slow-diffusion drug delivery vehicles for protein-type drugs, including erythropoietin, tissue-type plasminogen activator, synthetic haemoglobin and insulin. They can be used in applications, such as separation matrices (e.g. dialysis membranes). Collagen may be combined with the peptides to produce membranes for use as artificial skin. The MM may be used for making very thin, transparent fabric. Drugs which inhibit the self assembly of the peptides into filaments or filamentous membranes may be useful for treating Alzheimer's disease or scrapie infection. As they are resistant to proteolytic digestion and alkaline and acidic pH (such as stomach acid), drug delivery vehicles made of the MM could be taken orally. The charged residues and conformation of the MM promote cell adhesion and migration. The permeability of the MM also permits diffusion of small molecules, to the underside of cell monolayers, useful for tissue culture of differentiated cells and/or stratified cell layers.

SQ Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 other;

Query Match	10.4%;	Score 51.8;	DB 18;	Length 3083;
Best Local Similarity	55.7%;	Pred. No. 3.1e-05;		
Matches 123; Conservative	0;	Mismatches 92;	Indels 6;	Gaps 1

Qy	276	ACACTGATCCCCAAGACAGCTGGAGAACCAAGAATCATATTTGCAAGTCTTGTGACATTTGGCAAT	335
Db	1550	ACGCATGATGCTAGAGAGACTGGAAAACTGGCCGCAATTTGTTAGCTCATAGGGTTTGTCTAAG	1608
Qy	336	GTCAGATATCAAGGCTTACACAGACAGACAGATCAAAAGCGCTCAATAAGCAATGGTTTAAAA	395
Db	1610	TTCGCTTTAGAGCTACACGAAAGTCAATATCTATCAAGGCTCACACAAAAACAAGTTTCAAG	1668
Qy	396	CATCACCCAGAACAAAGGAAAAAGCGCTGGTGGTGAACCAATTAAGAAAGAGATATATGCTAC	455
Db	1670	TTCACATCTCCAGAACAAACATCTGCTGCTGGGGTGAATTGGACACAG-----ATGGCTTT	1723
Qy	456	TTCATCTTGCATACACTAAAGCTTTATGAATATGTATCTGATCC	496
Db	1724	TTCACAGATTTTCAAAAGGCTTGGAAACTTTCAGCTGATTC	1764

RESULT 8
AA211218
ID AA211218 standard; DNA; 3083 BP

AC	AAZ11218;
XX	
DT	09-NOV-1999 (first entry)
XX	
DE	Human zootin coding sequen

KW Membranforming peptide: cell culturing, macroscopic membrane;
 KW amphiphilic peptide: slow-diffusion drug delivery system; cell
 KW artificial skin; separation matrix; artificial tissue; scrapie infection,
 KW Alzheimer's disease, liver cirrhosis, kidney amyloidosis;
 KW protein conformational disease; human; zooin; ZO01; ss.

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 1292..2593
 FT /tag= a
 FT /product= zuotin
 XX
 XX USS955343-A.
 XX
 XX 21-SEP-1999.
 XX
 XX 22-AUG-1994; 94US-0293284.
 XX
 XX 22-AUG-1994; 94US-0293284.
 PR 28-DEC-1992; 92US-0973326.
 XX
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA
 XX
 PI Diersio CM, Holmes T, Lockshin C, Rich A, Zhang S;
 XX
 XX WPI: 1999-539576/45.
 DR P-PSDB: AAI32954.
 XX
 XX
 PT Cell cultures utilizing stable macroscopic membranes formed by the
 self-assembly of amphiphilic peptides
 XX
 PS Disclosure: Column 33-38; 49pp; English.
 CC
 CC This sequence encodes the human zuotin (ZUO1) protein. Fragments of the
 protein can be used in the in vitro method of the invention. The method
 is for culturing cells utilizing stable macroscopic membranes formed by
 the self-assembly of amphiphilic peptides. The peptides are salt-induced
 to form insoluble and protease-resistant protein filaments with a
 beta-sheet secondary structure. The membranes may be used in a
 wide variety of medical, research, industrial and biomaterial
 applications such as slow-diffusion drug delivery systems, artificial
 skin and separation matrices. The membranes may be used to support in
 vitro cell attachment and growth and for supporting artificial tissue
 (e.g. for in vivo use as implants). They are particularly useful as
 experimental models for Alzheimer's disease and scrapie infection and so
 may be used in disease modelling experiments and to assay for agents
 which modulate the disease processes. Additionally, they may be used in
 this way to study liver cirrhosis, kidney amyloidosis and other protein
 conformational diseases. The membrane in the membrane/cell mixture
 produced in the method: (i) supports cellular attachment and growth;
 CC (ii) self-assembles to form large, macroscopic membranes that are
 CC insoluble and stable in aqueous solutions, serum and ethanol; (iii) is
 CC highly resistant to heat, alkaline/acidic pHs, chemical denaturants and
 CC proteolytic digestion; (iv) is non-cytotoxic and non-immunogenic; (v) is
 CC visible to the naked eye if dyed but is otherwise transparent; (vi) may
 CC form thin, permeable, high tensile strength and fibrous structures with
 CC simple structures, high tensile strength and a porous structure;
 CC (vii) may be metabolised by humans and animals; (viii) is inexpensive to
 CC produce; and (ix) can be produced and stored in sterile conditions.
 XX
 XX Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 other;
 SO
 Query Match 10.4%; Score 51.8; DB 20; Length 3083;
 Best Local Similarity 55.7%; Pred. No. 3.1e-05;
 Matches 123; Conservative 0; Mismatches 92; Indels 6; Gaps 1;
 QY 276 ACACATGATGCCAAGACTGGAAGAACCAAGATCTTATGCAAGTCTTGAGCGCCAT 335
 DB 1550 ACTCATGATGCTAGAGACTGGAAGAACCTCCGATTTGCTGCTATGGGTTGCTAAG 1509
 QY 336 GTCAGATTAACAAGCTACACAGAGACAGATCAAGAGCTCTTAAGAGANGTTTAA 395
 DB 1610 TTGCGTTTACAGAGCTACGTAAGAGTCAATCATCAAGCTCACAGAAAACAAGTTGTCAG 1669
 QY 396 CACACACCAAGAAAGCGAAAGAGCTGTGTAACCAAAAAGAGAGAGATAGACTAC 455
 DB 1670 TACCATCCAGACAGCAATCTCTGCTGTGCTAGTTGACCAAG-----ATGGCTTT 1723

QY 456 TTCACCTTGCACTAATAAGCTATGAAGATATCTGATCC 496
 DB 1724 TTCAGATTTATTCAAAAGCGCTTTGAAGAACTTGACGATTC 1764
 RESULT 9
 ID ABL16646/c
 XX ABL16646 standard; DNA; 4791 BP.
 XX
 AC ABL16646;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 1411.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 interactions -
 XX
 PS Claim 1; SEQ ID NO 1411; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU57737-ABU72072),
 CC (ABU57737-ABU72072),
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 4791 BP; 1218 A; 1152 C; 1049 G; 1372 T; 0 other;
 SO
 Query Match 8.5%; Score 42.4; DB 23; Length 4791;
 Best Local Similarity 59.0%; Pred. No. 0.029;
 Matches 92; Conservative 0; Mismatches 61; Indels 3; Gaps 1;
 QY 337 TGAGATTAACAAGCTACACAGAGATCAAGACACTCATTAAGACATGTTTAAAC 396
 DB 3265 TTAGATTAAGAGCCAGGAGAGATGATGTTGACGGGCTTACAGGCCCATGCTGCTGC 3206
 QY 397 ATCAACCCAGACAAAGCGAAAGAGCTGTGTAACCAATTAAGAAAGAGATTAATGACTACT 456
 DB 3205 AACATCCCGATTAAGCGGAAAGCGGAGGAGAGAGTCAATCA--GGACGATGATTA 3149
 QY 457 TCACCTTGATTAAGAGCTTATGAATGTTATCTG 492
 DB 3148 TCACATGCATTAACCAAGCTTACAGATTAAGTGGGTG 3113
 RESULT 10


```

FT      from B. subtilis"
FT      39242..39904
FT      /tag= j
FT      /label= MG033
FT      /note= "Previously identified as MORF-20100, the
FT      encoded protein shows 35.90 percentage
FT      identity to glycerol uptake facilitator
FT      (glpF) from B. subtilis"
FT      CDS
FT      complement (39873..40514)
FT      /tag= k
FT      /label= MG034
FT      /note= "Previously identified as MORF-20101, the
FT      encoded protein shows 48.13 percentage
FT      identity to thymidylate kinase (tdk)
FT      from B. subtilis"
FT      CDS
FT      40543..41787
FT      /tag= l
FT      /label= MG035
FT      /note= "Previously identified as MORF-20102, the
FT      encoded protein shows 30.71 percentage
FT      identity to histidyl-tRNA synthetase (hiss)
FT      from Mycopacterium leprae"
FT      CDS
FT      complement (44751..46277)
FT      /tag= m
FT      /label= MG038
FT      /note= "Previously identified as MORF-20105, the
FT      encoded protein shows 46.83 percentage
FT      identity to glycerol kinase (glpK)
FT      from E. coli"
FT      CDS
FT      complement (46268..47422)
FT      /tag= n
FT      /label= MG039
FT      /note= "Previously identified as MORF-19831 and
FT      MORF-20106, the encoded protein shows 43.20
FT      percentage identity to glycerol-3-phosphate
FT      dehydrogenase (gdt2) from S. cerevisiae"
FT      CDS
FT      49377..49643
FT      /tag= o
FT      /label= MG041
FT      /note= "The encoded protein shows 48.86 percentage
FT      identity to phosphotransferase (ptsH) from Mycoplasma
FT      capricolum"
FT      CDS
FT      50060..51520
FT      /tag= p
FT      /label= MG042
FT      /note= "Previously identified as MORF-19832 and
FT      MORF-20108, the encoded protein shows 41.92
FT      percentage identity to spermidine/
FT      putrescine transport ATP-binding protein
FT      (potA) from E. coli"
FT      CDS
FT      51525..52382
FT      /tag= q
FT      /label= MG043
FT      /note= "Previously identified as MORF-20110, the
FT      encoded protein shows 26.51 percentage
FT      identity to spermidine/putrescine transport
FT      system permease protein (potB) from E. coli"
FT      CDS
FT      52366..53220
FT      /tag= r
FT      /label= MG044
FT      /note= "Previously identified as MORF-20111, the
FT      encoded protein shows 29.45 percentage
FT      identity to spermidine/putrescine transport
FT      system permease protein C (potC) from E. coli"
FT      CDS
FT      54658..55605
FT      /tag= s
FT      /label= MG046
FT      /note= "Previously identified as MORF-20112, the
FT      encoded protein shows 36.60 percentage
FT      identity to sialoglycoproteinase (gcp)
FT      from Pasteurella haemolytica"
FT      CDS
FT      complement (55670..58310)

```

```

FT      /tag= t
FT      /label= MG048
FT      /note= "Previously identified as MORF-19834,
FT      MORF-20114 and MORF-20115, the encoded protein
FT      shows 43.02 percentage identity to signal
FT      recognition particle protein (ifh) from B.
FT      subtilis"
FT      CDS
FT      58117..59079
FT      /tag= u
FT      /label= MG049
FT      /note= "Previously identified as MORF-20114 and
FT      MORF-20115, the encoded protein shows 44.78
FT      percentage identity to purine-nucleoside
FT      phosphorylase (decd) from E. coli"
FT      CDS
FT      59083..59754
FT      /tag= v
FT      /label= MG050
FT      /note= "Previously identified as MORF-20117, the
FT      encoded protein shows 83.03 percentage
FT      identity to deoxyribose-phosphate aldolase
FT      (decd) from Mycoplasma pneumoniae"
FT      CDS
FT      complement (64898..65731)
FT      /tag= w
FT      /label= MG056
FT      /note= "Previously identified as MORF-20122, the
FT      encoded protein shows 30.25 percent
FT      identity to the protein disclosed in
FT      GB:D26185.99 from B. subtilis"
FT      CDS
FT      complement (65713..66249)
FT      /tag= x
FT      /label= MG057
FT      /note= "Previously identified as MORF-20123, the
FT      encoded protein shows 36.90 percentage
FT      identity to the protein disclosed in
FT      GB:D26185.104 from B. subtilis"
FT      CDS
FT      81047..82597
FT      /tag= y
FT      /label= MG067
FT      /note= "Previously identified as MORF-19845, the
FT      encoded protein shows 28.84 percentage
FT      identity to glutamic acid specific protease
FT      (SPase) from Staphylococcus aureus"
FT      CDS
FT      91065..91919
FT      /tag= z
FT      /label= MG070
FT      /note= "Previously identified as MORF-20136, the
FT      encoded protein shows 34.8 percentage
FT      identity to ribosomal protein S2 (rps2)
FT      from Spirulina plantensis"

```

Query Match 8.5%; Score 42.4; DB 18; Length 580073;
Best Local Similarity 54.5%; Pred. No. 0.36;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```

OY 343 ACAAGCTACACAGACAGATCAAGCATTAAGCAATGTTTAAACATCACC 402
    |||||
DB 22434 AAAAGCGTAGTTCACAGACATTAAGCAATGTTTAAAGCTGCAATGCAATTCACC 22493
OY 403 CAGACAAAGGAAAGCAAGCGTGTGTAACCAATTAAGAGAGATGACTACTACT 462
    |||||
DB 22494 CCGATCGTCATTAAGCAAGCAAAATGAACTCAAAACAAATGAGAAAGATTAAAG 22553
OY 463 GCATACCTAAGCTTAAAGAAATGTTATCTGATCCAG 498
    |||||
DB 22554 AGTTAAATGAAGCATATGAAGTCTAAGTGATGAAG 22589

```

RESULT 12
AAK61812
ID AAK61812 standard; DNA: 1039 BP.
XX
XX AAK61812;

DT 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein coding sequence, t5-14.nt.

XX Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

OS Borrelia burgdorferi.

XX W09859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MED-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX P-PSDB; AAY20115.

XX Claim 1; Page 202; 275pp; English.

CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

XX Sequence 1039 BP; 526 A; 94 C; 212 G; 207 T; 0 other;

XX Query Match 8.2%; Score 41.2; DB 20; Length 1039;

XX Best Local Similarity 45.4%; Pred. No. 0.03; Mismatches 178; Indels 0; Gaps 0;

XX Matches 148; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

XX 124 AAGTGAACCTGTGGAGATGTTGACCTTTGTTAGAGAGAAACAGAAATGCTT 183

XX 45 AAGGAAATTAAGGATTTTATAGTTAGTTCAGCAAAAGATAAATAATCTT 104

XX 184 CTGCTCTTTTCAGAACTGGAGATTAAGAAAGTTATCCGAGAAATCAGAAATGAG 243

XX 105 CAAGTAGTTCAAAATGATGATTAATGCAAAATAATGCAAGAGATGAGATTAATG 164

XX 244 AATTCAGTTGAGAGATTTCCATGCTGAAACACTTGTATCCCAAGACTGGAAGACC 303

XX 165 AATTAATGAGGCGATGATCTTAATACAGAGCAATGACATGTTACAGATATTCGCG 224

XX 304 AAGATCATTTATGAGTTTGGATTTGGCACTGAGATGCAAGGCTACAGAGACAGA 363

XX 225 AAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 284

XX 364 TCAAGAGAGCTATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAGCAAGCTG 423

XX 285 AACAGAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 344

XX 424 GTGAACCAATTAAGAGAGATTAAT 449

XX 345 TAGAGAGAGAGAGAGAGAGAT 370

RESULT 13

AAK61811
ID AAK61811 standard; DNA; 1125 BP.

XX AAK61811;

XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein coding sequence, f5-14.nt.

XX Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

OS Borrelia burgdorferi.

XX W09859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MED-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX P-PSDB; AAY20114.

XX Claim 1; Page 202; 275pp; English.

CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

XX Sequence 1125 BP; 565 A; 101 C; 223 G; 236 T; 0 other;

XX Query Match 8.2%; Score 41.2; DB 20; Length 1125;

XX Best Local Similarity 45.4%; Pred. No. 0.032; Mismatches 178; Indels 0; Gaps 0;

XX Matches 148; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

XX 124 AAGTGAACCTGTGGAGATGTTGACCTTTGTTAGAGAGAAACAGAAATGCTT 183

XX 128 AAGGAAATTAAGGATTTTATAGTTAGTTCAGCAAAAGATAAATAATCTT 187

XX 184 CTGCTCTTTTCAGAACTGGAGATTAAGAAAGTTATCCGAGAAATCAGAAATGAG 243

XX 188 CAAGTAGTTCAAAATGATGATTAATGCAAAATAATGCAAGAGATGAGATTAATG 247

XX 244 AATTCAGTTGAGAGATTTCCATGCTGAAACACTTGTATCCCAAGACTGGAAGACC 303

XX 248 AATTAATGAGGCGATGATCTTAATACAGAGCAATGACATGTTACAGATATTCGCG 307

XX 304 AAGATCATTTATGAGTTTGGATTTGGCACTGAGATGCAAGGCTACAGAGACAGA 363

XX 308 AAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 367

XX 364 TCAAGAGAGCTATTAAGCAATGTTTAAACATCAACCAAGCAAGCAAGCAAGCTG 423

XX 368 AACAGAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 427

XX 424 GTGAACCAATTAAGAGAGATTAAT 449

Db 3937920 TCGCCGCCGTTGCCGCCGTTGCCGCCATCGCCGCCAGTGTGCCGCCAAG 3937861

RESULT 2

US-09-128-155-16
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc-feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match 15.1%; Score 75.6; DB 3; Length 152331;
Best Local Similarity 50.1%; Pred. No. 3.2e-05;
Matches 203; Conservative 0; Mismatches 195; Indels 7; Gaps 1;

DB 88 GGGGGTGGGGGGGCTCATGAACTCTTCCCGCCAGGCTAGGCTAGCCGAGAACCCCT 147
DB 21846 GGGTGTGTGGGGTGGGNGSTNNNACNNNNNNCCCGCCCGCCCGCCCGCCCT 21905
QY 148 GCAAGTGGGGGCTGAGCTCACTCCCGCTTCTTTCCTCCAGAGCTGGCCGCAATCC 207
DB 21906 CCCCCCGCCCGGAGGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCGCGCC 21965
QY 208 ATTCCAGGCGCGCGCGCGCGCCCGCCCGCCAGCGCCCGCGCGCGCGCGCGCA 267
DB 21966 CCCCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 22025
QY 268 GGTGACTGTCTCCCGGCTGCGCGCGCGCCCGAGAGTCCGCGGCTGTGCGCGCGC 327
DB 22026 GCG-----CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 22078
QY 328 TCGCGCGCTGTGCT 387
DB 22079 CCCCCCG 22138
QY 388 GGAACAGCGCGCGCTGAAAGCAGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 447
DB 22139 CG 22198
QY 448 GCG 492
DB 22199 CCG 22243

RESULT 3

US-09-050-863-2/c
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,863
FILING DATE: 30-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 949-8711
TELEFAX: (415) 949-8711

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 2580 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 14.9%; Score 74.4; DB 3; Length 2580;
Best Local Similarity 48.7%; Pred. No. 6.4e-05;
Matches 230; Conservative 0; Mismatches 241; Indels 1; Gaps 1;

QY 27 TCCCGGTGTTCTTCTGACAGCTGTGCGCCCGCCCTTCCCGCTGCTGGCTGAGATCC 86
DB 1358 TCCCTCTCTGCGCCCTGCTGCTGCTGCGCCCGCCCGCCCGCCCGCCCGCCCG 1299
QY 87 GGGGGGTGGGCGCTCATGAACTCTTCCCGCCAGCTCAAGGCTCAGCGCGCAACCCC 146
DB 1298 TCCCTCTGCGCCCTGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1239
QY 147 TCGAGTGGGGGCTGAGCTCAGTCCGCTTCTTTCCTCCAGAGCTGCGCGCGCA 206
DB 1238 TCCCTCTGCGCCCTGCTGCTGCTGCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1179
QY 207 CATTCAGAGCG 266
DB 1178 CCGCTCTGCTGCGCCCTGCGCCCGCCCGCCCGCGCGCGCGCGCGCGCGCGCG 1119
QY 267 AGTGGACTGTCTCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 326
DB 1118 TCGCGCTCTGCTGCTGCGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1060
QY 327 CCGCGCGCGCTGCG 386
DB 1059 CTCGTGCG 1000
QY 387 TGAACAGAGCGCGCTGAAGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 446
DB 999 CTCGTGCTGCG 940
QY 447 CCGCGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 498
DB 939 CTCGTGCG 888

RESULT 4

US-09-359-081-2/c
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:

```

OY      447 CGCGGCGGGCCAGGACGCGGCCCGCCCCGCTCAGCGGCACATTCAGTCCGCGGCG 498
         ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       939 CTCCTGCGCCCTCTGCGCCCTCTCTCTGTGCTCGCGCCCTCTCTCTGCTCTGCG 888

RESULT 5
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaji, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID903U1
; CURRENT APPLICATION NUMBER: US/09/130.114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: YEENBA
US-09-130-114-1

Query Match          14.9%; Score 74.4; DB 2; Length 5452;
Best Local Similarity 48.7%; Pred. No. 6.1e-05;
Matches 230; Conservative 0; Mismatches 241; Indels 1; Gaps 1;

OY      27 TCCCGGTTTCTTGTACAGCTGTGCGGGCCCCCCCCTTCCCTGTGTGAGCTGTGACTGCC 86
         ||| ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db     1447 TTCGCTCTCGCCCTCTGTCTGTCTGCGCCCTCTGTCTGTCTGTGCGCCCTCTGTCTGCC 1506

OY      87 GGGGGGTGGGGGCGCATGAATCTCTTCCGCGCACTTGAGGTTACAGGCCAGAACCCGC 146
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1507 TGCTTCTGCCCCCTCTGCGCCCTCTGTCTGTCTGTGCGCCCTCTGTCTGCCCTCTGTGCC 1566

OY     147 TGCAGGTGCGGGGTGAGCTGTGACGTCCGCGCTTCTTGGCTCCAGAGGCTGCGCCAGATGTC 206
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1567 TCTGCCCCCTCTCTCTGTCTGTCTGTGCGCCCTCTGTCTGTCTGTCTGTCTGTGCGCCCTCT 1626

OY     207 CATTCAGGCGCGCGCGCGCGCGCCCGCGCCAGCCAGCCAGCCCGCGCGCGCGAGCCCTTCC 266
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1627 CCGCTGTGCTCTGTGCCCCCTCTGTGCCCCCTCTGTCTGTCTGTGCGCCCTCTGTCTGCCCTCT 1686

OY     267 AGGTGGAGCTGTACCAGGTGTGCGCGCGCGCGCCAGAGATTCGCGCGCGCGGCTGTGCGGCGG 326
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1687 TGCCCCCTCTCTCTGTCTGTCTGTGCGCCCTCTGTCTGTCTGTCTGTCTGTCTGTGCGCCCTCT 1745

OY     327 CTGCGCGCGCTGTGCGCGCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCTGTACGG 386
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1746 CTCCTGTGCCCCCTCTCTGTCTGTCTGTGCGCCCTCTGTCTGTCTGTGCGCGCGCGCTGTGCGCC 1805

OY     387 TGGACACAGCGGCCCTGAAGAAGCACTTCGCGGCGCGCCCTCTGCGCACCCCGCGCATGTGCG 446
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1806 CTCCTGTCTCTGCCCCCTCTCTGTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTG 1865

OY     447 CGCCCGGCGCGAGGCGCGCGCGCGCGCGCGCTCGCGCGCATATGCGCGCGCGC 498
         ||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db     1866 CTCCTGCGCCCTCTGTGCGCCCTCTCTGTCTGTGCGCGCGCGCGCGCGCGCGCTGTCTGTGCG 1917

RESULT 6
US-08-910-647-1/c
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; POLYNUCLEOTIDE DELIVERY
; NUMBER OF SEQUENCES: 4
```



```

QY 27 TCCCGGTGTTCTTGCAGACGTCGAGCCCGCCCTTCCCGCTGCTGAGGCTGACATCC 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2938 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 GGGGGGTGGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2878 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 147 TGGAGGTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2818 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 207 CATTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2758 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 AGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2698 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 CTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2639 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 387 TGGACACAGCGCGCTGAAAGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2579 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 447 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2519 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12

PCT-US93-04648-15/c
Sequence 15, Application PC/TUS9304648

GENERAL INFORMATION:

APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Natalie A., Mark, Melanie H.
TITLE OF INVENTION: HEPATOCTYME GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patlin (genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/04648

FILING DATE: 19930517

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/885971

FILING DATE: 18-MAY-92

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755,779P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881

TELEPHONE: 415/225-3216

TELEFAX: 415/952-9881

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 10596 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US93-04648-15

Query Match 14.9%; Score 74.4; DB 5; Length 10596;
Best Local Similarity 48.7%; Pred. No. 5.9e-05;
Matches 230; Conservative 0; Mismatches 241; Indels 1; Gaps 1;

```

QY 27 TCCCGGTGTTCTTGCAGACGTCGAGCCCGCCCTTCCCGCTGCTGAGGCTGACATCC 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2938 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 GGGGGGTGGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2878 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 147 TGGAGGTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2818 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 207 CATTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2758 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 AGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2698 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 CTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2639 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 387 TGGACACAGCGCGCTGAAAGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2579 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 447 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2519 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

US-09-056-556-182/c
Sequence 182, Application US/09056556

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556

FILING DATE: 07-Apr-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Marki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457

TELEPHONE: 415/225-3216

TELEFAX: 415/952-9881

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

Page 8

100

1

vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 126 a 371 c 266 g 182 t 9 others
ORIGIN

Query Match 63.3% Score 316.6 DB 9: Length 954;
Best Local Similarity 94.2% Pred. No. 1.8e-30;
Matches 325; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

QY 2 GGCCTACGCGGGGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 61
DB 103 GGGCCGCTGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 162
QY 62 TTCCCGCTGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 121
DB 163 TTCCCGCTGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 222
QY 122 CCTCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 181
DB 223 CCTCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 282
QY 182 GCTTCCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 241
DB 283 GCTTCCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 342
QY 242 CAGGCGCGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 301
DB 343 CAGGCGCGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 402
QY 302 GAGTCCGCGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 346
DB 403 GAGTCCGCGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 447

RESULT 4
AL561632 921 bp mRNA linear EST 16-FEB-2001
LOCUS AL561632 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0D008YEO9 5
DEFINITION prime, mRNA sequence.
ACCESSION AL561632
VERSION AL561632.1 GI:12909254
KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 921)
Li, W. B., Gruber, C., Jassse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 921
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D008YEO9"
/clone_1b="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6; Site: 1; NotI; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com>
BASE COUNT 105 a 375 c 286 g 141 t 14 others
ORIGIN

Query Match 61.8% Score 309.2; DB 9: Length 921;
Best Local Similarity 88.0% Pred. No. 1.5e-29;
Matches 331; Conservative 5; Mismatches 40; Indels 0; Gaps 0;

QY 2 GGCCTACGCGGGGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 61
DB 276 GGGCCGCTGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 335
QY 62 TTCCCGCTGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 121
DB 336 TTCCCGCTGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 395
QY 122 CCTCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 181
DB 396 CCTCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 455
QY 182 GCTTCCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 241
DB 456 GCTTCCAGGCTCAGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 515
QY 242 CAGGCGCGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 301
DB 516 CAGGCGCGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 575
QY 302 GAGTCCGCGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 361
DB 576 GAGTCCGCGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 635
QY 362 CCGGCGCGGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 377
DB 636 TCCCTCCTGCGCGCGAGGCGCATGTCCTGCGAGCGCTGCGCCGCC 651

RESULT 5
BE514418 902 bp mRNA linear EST 07-AUG-2000
LOCUS BE514418 601315642P1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3634139 5',
DEFINITION mRNA sequence.
ACCESSION BE514418
VERSION BE514418.1 GI:9721631
KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 902)
NIH-MGC <http://mgi.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapubs@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: image.llnl.gov
Plate: LCM338 row: d column: 12
High quality sequence start: 33
High quality sequence stop: 344.
Location/Qualifiers

FEATURES
source
1. 902
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3634139"
/clone_1b="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"


```

/lab_host="PH10B (phage-resistant)"
/notc-Organ: Lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAACGAC(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      149 a      315 c      266 g      171 t
ORIGIN

```


Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalobosbcm.tmc.edu
Villalobos, D.R., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Munzy, D.M., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/BLND at: <http://image.llnl.gov>
Series: IRAC Plate: 6 Row: K Column: 1
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
1. 2152
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3154686"
/issue_type="Mammary tumor. Brcal-/fl: MMTV-Cre model. 10 months old, gross tissue."
/clone_1lb="NCL_CGAP_Mam3"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
BASE COUNT 435 a 692 c 619 g 406 t
ORIGIN

Query Match 51.48; Score 256.8; DB 11; Length 2152;
Best Local Similarity 87.08; Pred. No. 3.5e-23;
Matches 282; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 125 CAGGGTCACGCCAGAACCCCTCAGAGTCGGGGCTGAGTCAGTCCCGCTTTTGGC 184
DB 67 CAGGGTCACGCCAGAACCCCTCAGAGTCGGGGCTGAGTCAGTCCCGCTTTTGGC 126
QY 185 TCCCAAGGCTGGCCAGAGTCATTCAGAGCCGCGCGCCGCCAGGCCAG 244
DB 127 TCCCAAGGCTGGCCAGAGTCATTCAGAGCCGCGCGCCGCCAGGCCAG 186
QY 245 GCCCGGCGGCGAGCCCTCAGAGTCGAGTCCTCCGCTGCTCGCCGCCAGGCCAG 304
DB 187 GCCCGGCGGCGAGCCCTCAGAGTCGAGTCCTCCGCTGCTCGCCGCCAGGCCAG 246
QY 305 TCCCGGCGGCGAGCCCTCAGAGTCGAGTCCTCCGCTGCTCGCCGCCAGGCCAG 364
DB 247 TCCCGGCGGCGAGCCCTCAGAGTCGAGTCCTCCGCTGCTCGCCGCCAGGCCAG 306
QY 365 GCCCTGCGGCGGCTCTACAGTGAGACAGAGCCCTGAGAGCTTCGGCGCCCT 424
DB 307 GCCCTGCGGCGGCTCTACAGTGAGACAGAGCCCTGAGAGCTTCGGCGCCCT 366
QY 425 CCGCAGCCCGCCAGTGTGCGG 448
DB 367 CAGCCGCGGTCGCGCGCGCTGCGG 390

RESULT 8

BE988668 379 bp mRNA linear EST 05-OCT-2000
LOCUS BE988668/c
DEFINITION UI-M-CG0P-bhy-g-11-0-UI-s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
ACCESSION BE988668
VERSION BE988668.1 GI:10665238
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Bonaldi, M.F., Lennon, G., and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H

FEATURES

Location/Qualifiers
1. 379
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0P-bhy-g-11-0-UI"
/clone_1lb="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pTZ19-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu.
TAG_Seq=None found"
BASE COUNT 46 a 109 c 181 g 43 t
ORIGIN

Query Match 49.68; Score 248.2; DB 10; Length 379;
Best Local Similarity 87.78; Pred. No. 5e-22;
Matches 271; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 174 GCGTTCCTCCCGCAGAGGCTGCGCCAGAGTCCATTCACAGCGCGCGCGCCCGC 233
DB 309 GCGTTCCTCCCGCAGAGGCTGCGCCAGAGTCCATTCACAGCGCGCGCGCCCGC 250
QY 234 CCAAGCGCCAGAGGCTGCGCCAGAGGCTGCGCCAGAGTCCATTCACAGCGCGCGCGCCCGC 293
DB 249 CCAAGCGCCAGAGGCTGCGCCAGAGGCTGCGCCAGAGTCCATTCACAGCGCGCGCGCCCGC 190
QY 294 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 353
DB 189 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 130
QY 354 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 413
DB 129 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 70
QY 414 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 473
DB 69 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 10
QY 474 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 482
DB 9 CCGCAGAGGTCGCGCGGCTGCTGCGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1

RESULT 9
BG773423 465 bp mRNA linear EST 15-MAY-2001
LOCUS BG773423
DEFINITION 602720325F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837135 5',
ACCESSION BG773423
VERSION BG773423.1 GI:14084076

C	10	74.4	14.9	799	19	AAV55831	Nucleotide sequenc
C	11	74.4	14.9	1926	21	AAAS0234	Epstein Barr virus
C	12	74.4	14.9	1926	22	AAP62902	EBV tethering prot
C	13	74.4	14.9	2580	21	AAAT5454	Nucleotide sequenc
C	14	74.4	14.9	5452	20	AAAG0923	Anti-sense strand
C	15	74.4	14.9	8705	20	AAZ33778	Vector plasmid DN
C	16	74.4	14.9	9600	19	AAV21663	Vector plasmid PCM
C	17	74.4	14.9	10380	20	AAAZ2248	Nucleotide sequenc
C	18	74.4	14.9	10596	14	AAOS1731	Plasmid pCISEBON f
C	19	74.4	14.9	10596	17	AAAT40348	Plasmid pCISEBON f
C	20	74.4	14.9	10596	20	AAAI5650	Nucleotide sequenc
C	21	74.4	14.9	16080	21	AAAS9553	DNA clone pCEK Cl.
C	22	73.2	14.6	985	19	AAV64548	M. tuberculosis im
C	23	73.2	14.6	985	19	AAAV4439	Mycobacterium tube
C	24	73.2	14.6	985	20	AAAT19349	M. tuberculosis an
C	25	73.2	14.6	985	20	AAAZ19137	M. tuberculosis re
C	26	72.6	14.5	320	21	AAA38183	Primer used in the
C	27	72.6	14.5	320	21	AAA38186	Primer used in the
C	28	72.6	14.5	1000	21	AAAG2484	Human colon cancer
C	29	72	14.4	114955	20	AAAS3491	Human adenosine A1
C	30	71.2	14.2	320	21	AAA38185	Primer used in the
C	31	70.8	14.2	2668	22	AAAT5409	Human immune/haema
C	32	70.6	14.1	318	21	AAA38184	Primer used in the
C	33	70.2	14.0	1028	13	AAQ27092	Xy26 probe. Homo
C	34	70.2	14.0	1028	22	AAAG1579	1.0kb PstI fragmen
C	35	70.2	14.0	1028	22	AAAS01347	Human fragile X SY
C	36	69.8	14.0	1102	20	AAAT6933	Human gene express
C	37	69.2	13.8	4411529	22	AAI99682	Mycobacterium tubec
C	38	69	13.8	8438	15	AAOJ3500	DNA encoding pseud
C	39	68.8	13.8	955	22	AAAO1034	Sugarcane plant ge
C	40	68.8	13.8	1007	22	AAAS01022	Sugarcane plant ge
C	41	68.8	13.8	1698	23	AAAS84408	DNA encoding novel
C	42	68.6	13.8	2049	23	AAAT6210	DNA encoding novel
C	43	68.6	13.7	12001	16	AAAT6213	HSV I/ST region.
C	44	68	13.6	795	19	AAVS5830	FtCA insert stabili
C	45	68	13.6	2188	20	AAAT77506	Human ovarian tumo

ALIGNMENTS

RESULT 1						
ID	AAAD1114	strand:	DNA;	2389	Bp.	
XX	AAAD1114;					
XX						
DT	24-SEP-2001	(first entry)				
DE		Human small cell lung cancer associated gene, MAZ.				
XX						
KW		Human; small cell lung cancer; therapy; hCAP; nucleic acid; NA;				
KW		melanoma; cancer; colon; breast; head; neck; transitional cancer;				
KW		leiomyosarcoma; myc-associated zinc-finger protein; MAZ; cytosolic;				
KW		synovial sarcoma; ds.				
XX						
OS	Homo sapiens.					
PN	WO200153349-A2.					
XX						
PD	26-JUL-2001.					
XX						
FE	19-JAN-2001; 2001WO-US02015.					
XX						
PR	21-JAN-2000; 2000US-0489101.					
XX						
PA	(LUDW-) LUDWIG INST CANCER RES.					
PA	(SLOK) SLOAN KETTERING INST CANCER RES.					
PA	(CORR) CORNELL RES FOUND INC.					
PI	Stocker E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;					
XR	WPT; 2001-457597/49.					

xx Isolated polypeptide, used to treat or prognose a disorder
 PT characterized by expression of a hCAP e.g. cancer, is encoded by an
 PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
 xx
 PS Claim 57; Page 92-94; 152pp; English.

xx The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterised by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterised by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterised by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are neck, or non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
 CC The present sequence is a DNA encoding human myc-associated zinc-finger
 CC protein MZ. This small cell lung cancer associated gene is designated
 CC as NY-SCLC-5.
 xx

xx Sequence 2389 BP; 419 A; 805 C; 731 G; 434 T; 0 other;

Query Match 100.0%; Score 500; DB 22; Length 2389;
 Best Local Similarity 100.0%; Pred. No. 1.5e-63;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGGCTACAGGGGGGCGAGGCGATTTCCCGGTGTTTCTTTCACAGCTGTCGCCCCC 60
 DB 1 CGGCTACAGGGGGGCGAGGCGATTTCCCGGTGTTTCTTTCACAGCTGTCGCCCCC 60
 OY 61 CTTCCCGGCTGGGCGGCTGAGACTCCGGGGGGTGGGGCCCATGAACTCTTCCCGCC 120
 DB 61 CTTCCCGGCTGGGCGGCTGAGACTCCGGGGGGTGGGGCCCATGAACTCTTCCCGCC 120
 OY 121 ACCTCAGAGTCAAGCCCAAGAACCCCTTCAGAGTGGGGCTGAGCTTCAGTCCCGCTT 180
 DB 121 ACCTCAGAGTCAAGCCCAAGAACCCCTTCAGAGTGGGGCTGAGCTTCAGTCCCGCTT 180
 OY 181 TGCCCTCCAGGGCTGGGCGGCTGAGTCCAGGCGGCGGGGGCCCGCCCAAGCC 240
 DB 181 TGCCCTCCAGGGCTGGGCGGCTGAGTCCAGGCGGCGGGGGCCCGCCCAAGCC 240
 OY 241 CCAGGCGGCGGCGGCGGAGCCCTTCAGAGTGGACTTGTCCCGGTGCTCGCCGCCCA 300
 DB 241 CCAGGCGGCGGCGGCGGAGCCCTTCAGAGTGGACTTGTCCCGGTGCTCGCCGCCCA 300
 OY 301 GGAGTCCGCGGCGGCTGCTGCGGGCGGCTGCGCGGCTGCGCGGCTGCGCGGCG 360
 DB 301 GGAGTCCGCGGCGGCTGCTGCGGGCGGCTGCGCGGCTGCGCGGCTGCGCGGCG 360
 OY 361 CCGGCGGCGGCGGCGGCTGCTGCGGGCGGCTGCGCGGCTGCGCGGCTGCGCGGCG 420
 DB 361 CCGGCGGCGGCGGCGGCTGCTGCGGGCGGCTGCGCGGCTGCGCGGCTGCGCGGCG 420
 OY 421 CCGTCCGCGGCGGCGGCGGAGTGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480
 DB 421 CCGTCCGCGGCGGCGGCGGAGTGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 480
 OY 481 CGCCACTATGCGCGGCGGCGG 500
 DB 481 CGCCACTATGCGCGGCGGCGG 500

RESULT 2
 AAH26500 standard; cDNA; 2561 BP.

AAH26500:
 12-NOV-2001 (first entry)

DE Rabbit low density lipoprotein binding protein 2 (LBP-2) cDNA.

xx Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
 KM atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis;
 KM vaccine; ss.

xx Oryctolagus cuniculus.

xx Key Location/Qualifiers

FT CDS 246..1928

FT /tag= a

PN W0200164874-A2.

PF 07-SEP-2001.

PR 28-FEB-2001; 2001WO-US06356.

PR 02-MAR-2000; 2000US-0517849.

PR 14-JUL-2000; 2000US-0616289.

PR (BOST-) BOSTON HEART FOUND INC.

PI Lees AM, Lees RS, Law SW, Arjona AA;

PI WPI: 2001-565505/63.

PI P-PSDB; AAB82807.

PT New isolated low density lipoprotein binding polypeptide for treating,

PT diagnosing and/or identifying therapeutic agents for atherosclerosis

PS Claim 4; Fig 2A; 143pp; English.

xx The present sequence is that of cDNA encoding novel rabbit
 CC low density lipoprotein binding protein 1 (LBP-2, see AAB82807).
 CC The cDNA was isolated following screening of a rabbit cDNA library
 CC for clones encoding LBPs that bound to both native low density
 CC lipoprotein (LDL) and methyl LDL. The invention provides claimed
 CC polynucleotides encoding novel polypeptides which are capable of
 CC binding to native and methylated LDL, the isolated polypeptides,
 CC termed LBPs, and biologically active fragments and analogues of
 CC them, as well as expression vectors, cells and methods of producing
 CC the LBPs. Also claimed are methods of determining if an animal is
 CC at risk for atherosclerosis, methods for evaluating an agent for
 CC use in treating atherosclerosis, and methods for treating a cell
 CC having an abnormality in structure or metabolism of LBP.
 CC Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed.

xx Sequence 2561 BP; 372 A; 937 G; 879 G; 373 T; 0 other;

Query Match 18.1%; Score 90.6; DB 22; Length 2561;
 Best Local Similarity 54.8%; Pred. No. 2.7e-05;
 Matches 200; Conservative 0; Mismatches 164; Indels 1; Gaps 1;

OY 133 CGCCGAGAACCCCTGCGAGTGGGGCTGAGTCCAGTCCCGCTTTCCTCCAGAG 192
 DB 614 CGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 673
 OY 193 CTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 252
 DB 674 CGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 733
 OY 253 GCGCGAGCCCTCCAGGTGAGTCTCCCGGTGCTGCGGCGGCGGCGGCGGCGGCG 312
 DB 734 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 793
 OY 313 GCGTGCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 372
 DB 794 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 853
 OY 373 CGCCGCTCTAGGCTGAGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 432

CC chronic myelogenous leukaemia, rheumatoid arthritis, psoriasis and

DR WPL; 1998-312463/21.
XX
XX

OY 387 TGGACACAGCGGCGCTGAAGCAGCCTCGCGCGCCCGCCGACCGCCGCACTGTGCG 446
DB 1806 CTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1865
OY 447 GCGCCGCGCGCGAGCG 498
DB 1866 CTCCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1917

RESULT 15

AAZ23778
ID AAZ23778 standard; DNA; 8705 BP.

AC AAZ23778;

DT 14-JAN-2000 (first entry)

DE Vector pshuttle DNA.

KW Antisense; DNA library; identification; multiple cloning site; MCS;
inhibition; ss.

OS Synthetic.

PN WO950457-A1.

PD 07-OCT-1999.

PF 28-MAR-1999; 99WO-US06742.

PR 28-MAR-1998; 98US-0079792.

PR 06-NOV-1998; 98US-0107504.

PA (UTAH) UNIV UTAH RES FOUND.

PI Ruffner DE, Pierce ML, Chen Z;

DR WPI; 1999-610866/52.

PT Production of antisense libraries, used for identifying antisense
agents and for identifying target sites for antisense-mediated
inhibition of a selected gene -

PS Claim 16; Page 43-50; 63pp; English.

CC This invention describes a novel method for generating an antisense
CC library targeted to a selected RNA transcript. The methods can be used
CC for identifying antisense agents and for identifying target sites for
CC antisense-mediated inhibition of a selected gene. The use of a direct
CC library for target site selection significantly simplifies the screening
CC process, since only very small libraries need be prepared and assayed.
CC This sequence represents the vector pshuttle which is used in the method
CC of the invention.

SO Sequence 8705 BP; 1875 A; 2575 C; 2078 G; 2177 T; 0 other;

Query Match 14.9%; Score 74.4; DB 20; Length 8705;

Best Local Similarity 48.7%; Pred. No. 0.0043;

Matches 230; Conservative 0; Mismatches 241; Indels 1; Gaps 1;

OY 27 TCCCGGTTTCTTCTGACGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 86

DB 7312 TCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7371

OY 87 GGGGGGTGGGCGGCTGATGAATCTCTTCCCGGCACTCAGGTCAGCGCCAGAAACCCC 146

DB 7372 TGTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7431

OY 147 TGCAGGTGCGGCGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 206

DB 7432 TCTTGGCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7491

OY 207 CATTCAGGCG 266

DB 7492 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7551
OY 267 AGGTGACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 326
DB 7552 TGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7610
OY 327 CTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
DB 7611 CTCCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7670
OY 387 TGGACACAGCGGCGCTGAAGCAGCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 446
DB 7671 CTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7730
OY 447 GCGCCGCGCGAGCG 498
DB 7731 CTCCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7782

Search completed: October 10, 2002, 15:12:59
Job time : 1204 secs

FILING DATE: 26-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,859
FILING DATE: 29-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/621,443
FILING DATE: 29-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/621,457
FILING DATE: 30-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: G15181A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4405 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 408..2789
US-08-745-880-3

Query Match
Best Local Similarity 8.7%; Score 43.4; DB 2; Length 4405;
Matches 80; Conservativity 0; Mismatches 61; Indels 0; Gaps 0;

QY 334 GCGCCACGCGGCTCGCTCGGCTAGAGCGGCGCGCGCGCGTGAAGCGGTGAGCCCG 393
DB 209 GCGCGCGGCTCGGCG 150
QY 394 GTGCGCCCTCGGCG 453
DB 149 GCGCGGCTCGGCG 90
QY 454 TGCGTGGCGCGACG 474
DB 89 GGCTCGGCG 69

RESULT 4
US-08-480-382-3/c
Sequence 3, Application US/08480382
Patent No. 5986079
GENERAL INFORMATION:
APPLICANT: Barr, Phillip J.
APPLICANT: Brake, Anthony J.
APPLICANT: Kaufman, Rhadai J.
APPLICANT: Tekamp-Olson, Patricia
APPLICANT: Wasley, Louise
APPLICANT: Wong, Polly A.
TITLE OF INVENTION: Expression of PACE in Host Cells and
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson & Howson
STREET: Spring House
CITY: Spring House
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/480,382
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/885,972
FILING DATE: 20-MAY-1992
APPLICATION NUMBER: US 07/621,092
FILING DATE: 26-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,859
FILING DATE: 29-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/621,443
FILING DATE: 29-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/621,457
FILING DATE: 30-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: G15181A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4405 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 408..2789
US-08-480-382-3

Query Match
Best Local Similarity 8.7%; Score 43.4; DB 2; Length 4405;
Matches 80; Conservativity 0; Mismatches 61; Indels 0; Gaps 0;

QY 334 GCGCCACGCGGCTCGCTCGGCTAGAGCGGCGCGCGCGCGTGAAGCGGTGAGCCCG 393
DB 209 GCGCGCGGCTCGGCG 150
QY 394 GTGCGCCCTCGGCG 453
DB 149 GCGCGGCTCGGCG 90
QY 454 TGCGTGGCGCGACG 474
DB 89 GGCTCGGCG 69

RESULT 5
US-08-486-343A-1
Sequence 1, Application US/08486343A
Patent No. 6071695
GENERAL INFORMATION:
APPLICANT: OKAYNAK, ENGIN
APPLICANT: OPPERMAN, HERMAN
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
TITLE OF INVENTION: MORPHOGENIC PROTEIN EXPRESSION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
STREET: INC.
CITY: HOPKINTON
STATE: MA
COUNTRY: USA
ZIP: 07148
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```

COMPUTER : IBM PC compatible
OPERATING SYSTEM : PC-DOS/MS-DOS
SOFTWARE : PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA :
APPLICATION NUMBER: US/08/486,343A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, Edmund R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-091CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)-248-7000
TELEFAX: (617)-248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17415 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: homo sapiens
US-08-486-343A-1

```

Query Match	8.7%;	Score 43.4;	DB 3;	Length 17415;
Best Local Similarity	48.6%;	Pred. No. 0.34;		
Matches 119;	Conservative	0;	Mismatches 126;	Indels 0;
			Gaps	0
QY	206	CCCAAGACGACCCCTCCCGTCATTTGTGTGGGCTCGGGAGTGTGCGGGTTC	CCCCAGCCGC	265
Db	2493	CCCAACGCGGTCGCGACGCCACCCGTCGCTCGGGGCCGCCCTCTCTGGTTCGGACCGA		2552
QY	266	CQGGCGCGAGGACAAAGGGAGCGAGCCGCGCGGACGGGGACCCCGAGCTTGGCTGCT		325
Db	2553	GTCGCGAGAGGGGACGAGGGCGGCTCGCATTCCTCAGCCGCGATCCCGCGAGCTGCCGCA		2612
QY	326	CCCTCGGCTCGCCCGCAACGGGTTTCGCTCGGCTGAGACCGAGGGCGCCGCCATGAAGCGG		385
Db	2613	GCGCTTAGCGCACCCCTGGGGCACTCAGTAAACATTTGTGTGAGCGCTCTAGAGGGAAATGAA		2672
QY	386	TGAACCCGCGTCCGCCCTCGGGCGCGACGCGCCGCTCGGAGCGGCGAGCGAGCGGTGG		445
Db	2673	TGAACCCCACTGGGCACAGCTGGGGGGAGGGCGGGCGAGAGGACAGTGGGAGGCCGCCGG		2732
QY	446	CGCTG 450		
Db	2733	CGCGG 2737		

```

1 RESULT 6
2 PCT-US95-07349-1
3 Sequence 1, Application PC/TUS9507349
4 GENERAL INFORMATION:
5 APPLICANT:
6 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
7 TITLE OF INVENTION: MORPHOGEN EXPRESSION
8 NUMBER OF SEQUENCES: 7
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
11 ADDRESSEE: INC.
12 STREET: 45 SOUTH STREET
13 CITY: HOPKINTON
14 STATE: MA
15 COUNTRY: USA
16 ZIP: 07148
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: PatentIn Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: PCT/US95/07349

```

1 FILING DATE:
2 CLASSIFICATION:
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 07/938,021
5 FILING DATE: 28-AUG-1992
6 ATTORNEY/AGENT INFORMATION:
7 NAME: KELLEY, ROBIN D
8 REGISTRATION NUMBER: 34,637
9 REFERENCE/DOCKET NUMBER: CRP-091PC
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: (508)-435-9001
12 TELEFAX: (508)-435-0992
13 INFORMATION FOR SEQ ID NO: 1:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 17415 base pairs
16 TYPE: nucleic acid
17 STRANDEDNESS: single
18 TOPOLOGY: linear
19 MOLECULE TYPE: DNA (genomic)
20 PCT-US93-07349-1

	Query Match	8.7%;	Score 43.4;	DB 5;	Length 17415;
	Best Local Similarity	48.6%;	Pred. No. 0.34;		
Matches 119;	Conservative	0;	Mismatches 126;	Indels	Gaps
OY	CCCGAGACGCCCTTCCGTCATTTMTTGSGGCGGGGAOTGTGCGGTGCCTCCGAGCGGC	265			
Db	2493 CCGACGCGGGTCCGAGACCACCGCTCMTGCTGGGGCCCGCTTCTGTGTCGGAGCCGCA	2552			
OY	CQGGCGCGGAGCAAAAGGAGCGAGCGGCGCGGAGCGGGAGCGGGAGCTTCMGCCG	325			
Db	2253 GTGGCGAGAGGCGAAGGGCGGGCTCCGATTCTCCAGCGGCATCCCCCGAGCTCCGSCA	2612			
OY	CCCTCGCTGCCCCCACGGGTTTCGTCGCTGTAAGCGCAGGGCGCGCGCATGAAGAAGCGG	385			
Db	2613 GGCCTTAGGACACCCCCTGGGCACACTAGTAACATTGTGTGACGGCTCTAGAGGGATTGAA	2672			
OY	TGAACCCCGGTGCGCCCTTGGGCGCCCAAGCGCCGTGGGCTGGCGCGCGCGGACTG	445			
Db	2673 TGAACCCCACTGGGCAAGCTGGGGGAGGCGCGGGCCAGGGCAAGTGGAGGCCCGCCCG	2732			
OY	446 CGCTG 450				
Db	2733 CGCGG 2737				

```

1  RESULT 7
2  US-09-320-878-21/c
3  Sequence 21, Application US/09320878A
4  Patent No. 6117659
5  GENERAL INFORMATION:
6  APPLICANT: ASHLEY, Gary
7  APPLICANT: BETLACH, Melanie C.
8  APPLICANT: BETLACH, Mary C.
9  APPLICANT: MCDANIEL, Robert
10 APPLICANT: TANG, Li
11 TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
12 FILE REFERENCE: 300622002120
13 CURRENT APPLICATION NUMBER: US/09/320,878A
14 CURRENT FILING DATE: 1999-05-27
15 EARLIER APPLICATION NUMBER: CIP OF 09/141,908
16 EARLIER FILING DATE: 1998-08-28
17 EARLIER APPLICATION NUMBER: CIP OF 09/073,538
18 EARLIER FILING DATE: 1998-05-06
19 EARLIER APPLICATION NUMBER: CIP OF 08/846,247
20 EARLIER FILING DATE: 1997-04-30
21 EARLIER APPLICATION NUMBER: 60/119,139
22 EARLIER FILING DATE: 1999-02-08
23 EARLIER APPLICATION NUMBER: 60/100,880
24 EARLIER FILING DATE: 1998-09-22
25 EARLIER APPLICATION NUMBER: 60/087,080
26 EARLIER FILING DATE: 1998-05-28
27 NUMBER OF SEQ ID NOS: 34

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 5970
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-876-21

```

Query Match	8.6%	Score 43.2;	DB 3;	Length 5970;
Best Local Similarity	47.4%	Pred. No. 0.36;		
Matches 129;	Conservative	0;	Mismatches 143;	Indels 0;
				Gaps 0;

QY	219	TCCCGTCAATTGTTGGGCTCGGAGATGTGTGCGTGGCCCGAGACGGCGGAGGAGGC	278
Db	5793	TCCCGGGGCTGTGTGTTCGCCCGGGCCACACGGGTGCACATCGGGCGTCCCGCGGTCC	5734
QY	279	AAAGGAGCGGATCCGGCGCGGACGGGGGCCGAGACTTCCTGGCTCCCTCGCTCGGCC	338
Db	5733	TGAGGTCCCGGTTCGCCACAGCGGTCCCTGACGTCTCCGTGTGCTCCGAGCGTCTGGCGGT	5674
QY	339	CAGCGGTTTCCTCTCGGTAGAGCGGACGGCGCGCGCGATGAAGCGGTGATCCCGGTGCG	398
Db	5673	CTCCGGGTGCGGACACGGAGCGCTCTGGGCGCCTCTGGGCGGTACAGAGATCCCATGGCA	5614
QY	399	CCCTCGGGCGCCAGAGCGCCCTCTGGGCTGCGGCGCGGGGAGCTGCGCTGTGCGCTGCT	458
Db	5613	CTTGCGCGCGCCCGGCTGTGCGACACCGGCGTGTCTCCCGGCGGCGCGGTGGCCCTGCGCGG	5554
QY	459	GGCCAGACACGCGCCACAGCCCTGGGTGGCTCG	490
Db	5553	GTCTGTGCGACGGGAGACACCGGGGTGGCTTAG	5522

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

; STRANDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..13040
;
US-08-658-136-4

```

Query Match	8.6%;	Score 43.2;	DB 3;	Length 14060;
Best Local Similarity	50.0%;	Pred. No. 0.37;		
Matches 108;	Conservative 0;	Mismatches 108;	Indels 0;	Gaps 0;

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

NAME/KEY: CDS
 LOCATION: 14046..20036
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 20110..31284
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 31329..36071
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 36155..41830
 US-08-804-227C-7

Query Match 8.6%; Score 43.2; DB 2; Length 44377;
 Best Local Similarity 48.0%; Pred. No. 0.4;
 Matches 123; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 245 GTTCGGGTGCCCCGAGCGCCGCGGCGAGGCAAGGAGCGCGCGCGAGC 304
 DB 43302 GTTGGGGGCGCATCCGCGGCTGAGCTGGGGCGCGAGCGAGTCCCGCGCGCG 43243
 QY 305 GGGCCCGAGCTTGGCTGCTCTCTGCTGCGCCCGAGCGGGGTGCGTGGAGAGCGCA 364
 DB 43242 TGGCGGGGTGCGGGCGCCGCGCATCCGAGAGCTGAGCATCAGCTCAGCGCGTAGAGGCC 43183
 QY 365 GGGCGCGCGCATGTAAGCGCGTGAACCCGCGCCCTCGGGCGCGCAAGCGCGTGGC 424
 DB 43182 GGGCGACCTTCCGCGCAAGGTGCAACCGCGGCTCAGCCGTGGGCAAGCGCGCGCA 43123
 QY 425 GCTGGCGGCGCGGAGCTGCGCTGCGCTGCGCGGAGCGGCGGCGGCGTGGAGT 484
 DB 43122 CTCGTGGGGGCGGTGGTGGCGCGCGCGCGAGGGCGCTGGGGCGGGGCGTACGGGCTGAGCA 43063
 QY 485 GCTCCGAGCGCGCGC 500
 DB 43062 CCGGGCGCGCGCGC 43047

RESULT 10
 US-08-804-198-1/c
 Sequence 1, Application US/08804198
 Patent No. 5945320
 GENERAL INFORMATION:
 APPLICANT: Burgett, Stanley G.
 APPLICANT: Kuhnloss, Stuart A.
 APPLICANT: Rao, Nagendra R.
 APPLICANT: Richardson, Mark A.
 APPLICANT: Rostek, Paul R., Jr.
 TITLE OF INVENTION: PLATEOLIDE SYNTHASE GENE
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PAUL R. CANTRELL 1138
 STREET: LILLY CORPORATE CENTER
 CITY: INDIANAPOLIS
 STATE: IN
 COUNTRY: USA
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: Macintosh 7.0
 SOFTWARE: Microsoft Word 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/804.198
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: CANTRELL, PAUL R.
 REGISTRATION NUMBER: 36,470
 REFERENCE/DOCKET NUMBER: P9113
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3885
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 44377 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 350..14002
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 14046..20036
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 20110..31284
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 31329..36071
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 36155..41830
 US-08-804-198-1

Query Match 8.6%; Score 43.2; DB 2; Length 44377;
 Best Local Similarity 48.0%; Pred. No. 0.4;
 Matches 123; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 245 GTTCGGGTGCCCCGAGCGCCGCGGCGAGGCAAGGAGCGCGCGCGAGC 304
 DB 43302 GTTGGGGGCGCATCCGCGGCTGAGCTGGGGCGCGAGCGGCGAGTCCCGCGCGCG 43243
 QY 305 GGGCCCGAGCTTGGCTGCTCTCTGCTGCGCCCGAGCGGGGTGCGTGGAGAGCGCA 364
 DB 43242 TGGCGGGGTGCGGGCGCCGCGCATCCGAGAGCTGAGCATCAGCTCAGCGCGTAGAGGCC 43183
 QY 365 GGGCGCGCGCATGTAAGCGCGTGAACCCGCGCCCTCGGGCGCGCAAGCGCGTGGC 424
 DB 43182 GGGCGACCTTCCGCGCAAGGTGCAACCGCGGCTCAGCCGTGGGCAAGCGCGCGCA 43123
 QY 425 GCTGGCGGCGCGGAGCTGCGCTGCGCTGCGCGGAGCGGCGGCGGCGTGGAGT 484
 DB 43122 CTCGTGGGGGCGGTGGTGGCGCGCGCGCGAGGGCGCTGGGGCGGGGCGTACGGGCTGAGCA 43063
 QY 485 GCTCCGAGCGCGCGC 500
 DB 43062 CCGGGCGCGCGCGC 43047

RESULT 11
 US-09-428-517-1/c
 Sequence 1, Application US/09428517
 Patent No. 6251636
 GENERAL INFORMATION:
 APPLICANT: Betlach, Mary C.
 APPLICANT: Shah, Sanjay Krishnakant
 APPLICANT: McDaniel, Robert
 APPLICANT: Tang, Li
 TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
 FILE REFERENCE: 30062-20029.00
 CURRENT APPLICATION NUMBER: US/09/428.517
 EARLIER FILING DATE: 1999-10-28
 EARLIER APPLICATION NUMBER: 60/120,254
 EARLIER FILING DATE: 1999-02-16
 EARLIER APPLICATION NUMBER: 60/106,100
 EARLIER FILING DATE: 1998-10-29
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 50937
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA

Db 3641 CCMACGATGCGCCCGCCGCGCCGCGCCGCGCGCTGCGCGCTGCGCTGCGCTGCG 3700
 QY 369 GCGCGGATGAAAGCGGATGAGCCGCGCTGCGCGCCGCGCAAGCGCGCTGCGGCTG 428
 Db 3701 GCGCGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGCGCTGCGCGCGCTG 3760
 QY 429 CCG 464
 Db 3761 CCG 3796

RESULT 14

US-09-128-155-16/c
 : Sequence 16, Application US/09128155
 : Patent No. 6117654
 : GENERAL INFORMATION:
 : APPLICANT: Pan, Yang
 : TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
 : FILE REFERENCE: 09404/052001
 : CURRENT APPLICATION NUMBER: US/09/128,155
 : EARLIER FILING DATE: 1998-08-03
 : EARLIER APPLICATION NUMBER: US 60/091,650
 : EARLIER FILING DATE: 1998-07-02
 : EARLIER APPLICATION NUMBER: US 60/054,646
 : EARLIER FILING DATE: 1997-08-04
 : NUMBER OF SEQ ID NOS: 18
 : SOFTWARE: FASTSEQ for Windows Version 3.0
 : SEQ ID NO 16
 : LENGTH: 152331
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)..(152331)
 : OTHER INFORMATION: n = A,T,C or G
 : US-09-128-155-16

Query Match 8.6%; Score 43; DB 3; Length 152331;
 Best Local Similarity 47.7%; Pred. No. 0.47;
 Matches 124; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 240 GGAAGTGTGCGGCG 299
 Db 22171 GGGGCG 22112
 QY 300 GGACGGGCG 359
 Db 22111 GGGGCG 22052
 QY 360 GCGCAGGCG 419
 Db 22051 GGGGCG 21992
 QY 420 GTCGCGCTGCG 479
 Db 21991 GCGGCG 21932
 QY 480 GAGTGGCTCG 499
 Db 21931 GCG 21912

RESULT 15

US-08-034-650-9
 : Sequence 9, Application US/08034650
 : Patent No. 5641671
 : GENERAL INFORMATION:
 : APPLICANT: BOS, Jannetje W.
 : APPLICANT: FRENKEN, Leon G.
 : APPLICANT: VERKERS, Cornelis T.
 : APPLICANT: VISSER, Christiaan
 : TITLE OF INVENTION: PRODUCTION OF ACTIVE PSEUDOMONAS GLUMAE

:/ TITLE OF INVENTION: LIPASE IN HOMOLOGOUS OR HETEROLOGOUS HOSTS
 : NUMBER OF SEQUENCES: 13
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
 : STREET: 1615 L. Street, N.W.
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : ZIP: 20036-5601
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/034,650
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/727,235
 : FILING DATE: 03-JUL-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Kokulis, Paul N.
 : REGISTRATION NUMBER: 16,773
 : REFERENCE/DOCKET INFORMATION:
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 861-3000
 : TELEFAX: (202) 822-0944
 : TELEX: 6714627 CUSH
 : INFORMATION FOR SEQ ID NO: 9:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2900 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 483..1556
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 1559..2617
 : US-08-034-650-9

Query Match 8.5%; Score 42.6; DB 1; Length 2900;
 Best Local Similarity 47.5%; Pred. No. 0.46;
 Matches 126; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 234 GCGTGGGAGTGTGCG 293
 Db 1561 GCGCGAGCGCGATCG 1620
 QY 294 GCGCGCGAGCG 353
 Db 1621 GTTCGGGCTGCG 1680
 QY 354 CGTAGAGCGCAGCG 413
 Db 1681 GCG 1740
 QY 414 GCGCGCGTGGCG 473
 Db 1741 GCG 1800
 QY 474 CAGCGTGGGCTCG 498
 Db 1801 CCGCGTGGGCTCG 1825

Search completed: October 10, 2002, 17:54:40
 Job time : 165.733 secs

```

/organism="Homo sapiens"
/db_xref="taxon.9606"
/clone="IMAGE:5285103"
/clone_1_id="NH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/organ="Organ: brain; Vector: pBluescriptpl (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcaca
g); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTNN-3',
Size selected for average insert size 2.5 kb and

```

normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHRI, National Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 135 a 268 c 320 g 119 t 1 others

Query Match 90.0%; Score 450.2; DB 10; Length 843;
Best Local Similarity 98.6%; Pred. No. 2.8e-66;
Matches 486; Conservative 0; Mismatches 3; Indels 4; Gaps 3;

10 GGGCGGTTGTGATGATACCGGAGTGGGATATCCCGGGCTAGGGAGCGGCGG-C 67
|||||
23 GGGCGGTTGTGATGATACCGGAGTGGGATATCCCGGGCTAGGGAGCGGCGGCG 82
|||||
68 CCGATGGGCTTGTGAGGAGTCCGAAGGAGTACTAGACACCGGGGGGCTACTTT 127
|||||
83 GCGATGGGCTTGTGAGGAGTCCGAAGGAGTACTAGACACCGGGGGGCTACTTT 142
|||||
128 TCTTCGGTGTGCTTTGCTTTTCTTTTCTTTGGGCTGGGCTGAGTGTCCGCACTGAGC 187
|||||
143 TCTTCGGTGTGCTTTGCTTTTCTTTTCTTTGGGCTGGGCTGAGTGTCCGCACTGAGC 201
|||||
188 AAGATTTCCCTGCTAAACCCAGAGGACCCCTCCGCAATTTGGGGCTGGGAGTGTG 247
|||||
202 AAGATTTCCCTGCTAAACCCAGAGGACCCCTCCGCAATTTG-TGGGCTGGGAGTGTG 260
|||||
248 GCGGTGCCCCGAGCGCGCGCGGCGGAGGCAAGGAGCGAGCGCGCGGCGGAGCGG 307
|||||
261 GCGGTGCCCCGAGCGCGCGCGGCGGAGGCAAGGAGCGAGCGCGCGGCGGAGCGG 320
|||||
308 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 367
|||||
321 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
|||||
368 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
|||||
381 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
|||||
428 GCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 487
|||||
441 GCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
|||||
488 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
|||||
501 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
|||||

RESULT 2
BI552992 570 bp mRNA linear EST 05-SEP-2001
LOCUS 603193628F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5264694 5',
DEFINITION mRNA sequence.
ACCESSION BI552992
VERSION BI552992.1 GI:15440304
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: egapbs-remail.nih.gov
Tissue Procurement: MIKLOS Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHRI), Shihaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: L14M1667 row: h column: 07
High quality sequence stop: 543.
Location/Qualifiers

FEATURES
source
1..570
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5264694"
/clone_1ib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 72 a 174 c 239 g 85 t

Query Match 88.8%; Score 444; DB 10; Length 570;
Best Local Similarity 98.3%; Pred. No. 3.3e-65;
Matches 470; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

10 GGGCGGTTGTGATGATACCGGAGTGGGATATCCCGGGCTAGGGAGCGGCGG-C 67
|||||
22 GCGCGGTTGTGATGATACCGGAGTGGGATATCCCGGGCTAGGGAGCGGCGGCGG 81
|||||
68 CCGATGGGCTTGTGAGGAGTCCGAAGGAGTACTAGACACCGGGGGGCTACTTT 127
|||||
82 GCGATGGGCTTGTGAGGAGTCCGAAGGAGTACTAGACACCGGGGGGCTACTTT 141
|||||
128 TCTTCGGTGTGCTTTGCTTTTCTTTTCTTTGGGCTGGGCTGAGTGTCCGCACTGAGC 187
|||||
142 TCTTCGGTGTGCTTTGCTTTTCTTTTCTTTGGGCTGGGCTGAGTGTCCGCACTGAGC 200
|||||
188 AAGATTTCCCTGCTAAACCCAGAGGACCCCTCCGCAATTTGGGGCTGGGAGTGTG 247
|||||
201 AAGATTTCCCTGCTAAACCCAGAGGACCCCTCCGCAATTTGAGGCTGGGAGTGTG 260
|||||
248 GCGGTGCCCCGAGCGCGCGGCGGAGGCAAGGAGCGAGCGCGCGGCGGAGCGG 307
|||||
261 GCGGTGCCCCGAGCGCGCGGCGGAGGCAAGGAGCGAGCGCGCGGCGGAGCGG 320
|||||
308 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 367
|||||
321 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
|||||
368 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
|||||
381 CCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
|||||
428 GCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 485
|||||
441 GCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
|||||

RESULT 3
BG337080 911 bp mRNA linear EST 27-FEB-2001
LOCUS 602434352F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4552357 5',
DEFINITION mRNA sequence.
ACCESSION BG337080
VERSION BG337080.1 GI:13143518
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against

Insert Length: 644 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 466.
Location/Qualifiers

FEATURES

source

1. 522

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1500639"

/clone_lib="NCI_CGAP_K1d3"

/lab_host="DH10B"

/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st
strand cDNA was primed with Not I and cloned into the Not
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 79 a 200 c 172 g 70 t 1 others
ORIGIN

Query Match 73.7%; Score 368.4; DB 9; Length 522;
Best Local Similarity 99.5%; Pred. No. 1.5e-52;

Matches 369; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 130 TTCGGGCTTTGCTTTTTCCTTTTTCCTTTGCGCTGCGCTAGCTGCGCCCACTAGCA 189

Db 522 TTCGGGCTTTGCTTTTTCCTTTTTCCTTTGCGCTGCGCTAGCTGCGCCCACTAGCA 463

OY 190 AGATTCCTCGTAACACAGAGACACCTCCGTCATTTGTTGGGCTCGGAGTGTCCG 249

Db 462 AGATTCCTCGTAACACAGAGACACCTCCGTCATTTGTTGGGCTCGGAGTGTCCG 403

OY 250 GGTGCCCCGAGCGCGCGCGCGGAGGCAAGGAGCGGACCGCGCGGAGCGGCGC 309

Db 402 GGTGCCCCGAGCGCGCGCGCGGAGGCAAGGAGCGGACCGCGCGGAGCGGCGC 343

OY 310 CGGACCTTGCCTCCCTCGCTGCGCCACAGGGGTTGCTCGCTAGAGCGCAGGCG 369

Db 342 CGGACCTTGCCTCCCTCGCTGCGCCACAGGGGTTGCTCGCTAGAGCGCAGGCG 283

OY 370 CGCGGATGAAGCGGAGCGCGCGCTGCGCCCTCGCGGCGGCAAGGCGCGCTCG 429

Db 282 CGCGGATGAAGCGGAGCGCGCGCTGCGCCCTCGCGGCGGCAAGGCGCGCTCG 223

OY 430 GCGCGGCGGAGCTGCGCTGCGCTGCGCGGAGCGGCGCAGCGCTGCGCTCG 489

Db 222 GCGCGGCGGAGCTGCGCTGCGCTGCGCGGAGCGGCGCAGCGCTGCGCTCG 163

OY 490 GCAGCGCGGCGC 500

Db 162 GCAGCGCGGCGC 152

RESULT 7 826 bp mRNA linear EST 05-SEP-2001

LOCUS B1546228 603188614F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5260155 5',

DEFINITION mRNA sequence.

ACCESSION B1546228 GI:15433540

VERSION B1546228.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 826)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)

DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
Plate: L14M11655 row: k column: 04
High quality sequence stop: 821.
Location/Qualifiers

FEATURES

source

1. 826

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5260155"

/clone_lib="NIH_MGC_95"

/tissue_type="hippocampus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 128 a 261 c 307 g 130 t
ORIGIN

Query Match 73.6%; Score 367.8; DB 10; Length 826;
Best Local Similarity 96.4%; Pred. No. 1.7e-52;

Matches 483; Conservative 0; Mismatches 7; Indels 11; Gaps 10;

OY 10 GGGGGGTTGTGAGTACCGGAGTGGGATCCCGGGCTAGGGAGCGCGGCG -C 67

Db 23 GGGGGGTTGTGAGTACCGGAGTGGGATCCCGGGCTAGGGAGCGCGGCGGCG 82

OY 68 CCGATCGGGCTTGTGAGTACCGGAGTGGGATCCCGGGCTAGGGAGCGCGGCTACTT 127

Db 83 CCGATCGGGCTTGTGAGTACCGGAGTGGGATCCCGGGCTAGGGAGCGCGGCTACTT 142

OY 128 TCTTCCGGTGTGCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTC 187

Db 143 TCTTCCGGTGTGCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTC 201

OY 188 AAAGATTTCCTGTAACACAGAGACACCTCCGTCATTTGTTGGGCTCGGAGTGTG 247

Db 202 AAAGATTTCCTGTAACACAGAGACACCTCCGTCATTTGTTGGGCTCGGAGTGTG 261

OY 248 GCGGTGCCCCGA -GCGGCGCGGCGCGGAGCAAGGAGCGGAGCGGCG -CGCGAGCG 305

Db 262 GCGGTGCCCCGA -GCGGCGCGGCGCGGAGCAAGGAGCGGAGCGGAGCGGCG 321

OY 306 GCGCGGAGCTTGCCTTCCCTCGCTGCGCCCA -GCGGGTTGCGTGGC -GTAGAGCG 363

Db 322 GCGCGGAGCTTGCCTTCCCTCGCTGCGCCCA -GCGGGTTGCGTGGC -GTAGAGCG 381

OY 364 AGGCGCG -GCGGAGTGAAGC -GGTGAAGCCGCTGCGCCCTCG -GGCGCAAGGCGCG 420

Db 382 AGGCGCG -GCGGAGTGAAGC -GGTGAAGCCGCTGCGCCCTCG -GGCGCAAGGCGCG 441

OY 421 TCGGCTGCGGCG -GCGGAGTGAAGC -GGTGAAGCCGCTGCGCCCTCG -GGCGCAAGGCGCG 479

Db 442 TCGGCTGCGGCG -GCGGAGTGAAGC -GGTGAAGCCGCTGCGCCCTCG -GGCGCAAGGCGCG 501

OY 480 GGTGCTCGGAGCGCGCGC 500

Db 502 GGGTGGCTCCGACGCCGCGC 522

RESULT 8
LOCUS BE327284 472 bp mRNA linear EST 14-JUL-2000
DEFINITION hw10a02.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3182474 3' similar to contains MER22.b3 TARI repetitive element ;, mRNA sequence.

ACCESSION BE327284
VERSION BE327284.1 GI:9201060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 472)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -400p from Glibcoo
High quality sequence stop: 421.
Location/Qualifiers
1..472
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3182474"
/clone_lib="NCI_CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker. Plasmid DNA from the normalized library NCI-CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 56 a 158 c 186 g 72 t

ORIGIN

Query Match 71.6%; Score 358; DB 10; Length 472;
Best Local Similarity 99.7%; Pred. No. 8.2e-51;
Matches 369; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 132 CCGGTGTTTGGC-TTTTTTTCCTTGGCTGGGCTGAGTGCCTCCACTGAGCAAA 190
|||||
Db 1 CCGGTGTTTGGCTTTTTTTTCTTGGCTGGGCTGAGTGCCTCCACTGAGCAAA 60
|||||

QY 191 GATTCCCTGTAAACCCAGAGCACCCTCCGTCATTTGTGGCTGGGAGTGTCCG 250
|||||
Db 61 GATTCCCTGTAAACCCAGAGCACCCTCCGTCATTTGTGGCTGGGAGTGTCCG 120
|||||

QY 251 GTTCCCGGAGCGCGCGGCGGAGGCAAGGAGCGGAGCCGCGGAGGCGCC 310
|||||
Db 121 GTTCCCGGAGCGCGCGGCGGAGGCAAGGAGCGGAGCCGCGGAGGCGCC 180
|||||

QY 311 GGAGCTTGCTGCTCCCTGCTGGCGGCCGAGCGGCTTCGCTCCGCTAGAGCGGAGCGC 370
|||||
Db 181 GGAGCTTGCTGCTCCCTGCTGGCGGCCGAGCGGCTTCGCTCCGCTAGAGCGGAGCGC 240
|||||

QY 371 GCGGATGAAGGCGGCTGAGCCCGGTGCGCCCTTCGCGGCCGCAAGCGCCGTGGCTGC 430
|||||
Db 241 GCGGATGAAGGCGGCTGAGCCCGGTGCGCCCTTCGCGGCCGCAAGCGCCGTGGCTGC 300
|||||

QY 431 GCGGCGGAGCAGCTGGCGCTGGCTCCCTGCGGACACAGCCACAGCCGTGGCTGC 490
|||||
Db 301 GCGGCGGAGCAGCTGGCGCTGGCTCCCTGCGGACACAGCCACAGCCGTGGCTGC 360
|||||

QY 491 CAGCCGCGGC 500
|||||
Db 361 CAGCCGCGGC 370
|||||

RESULT 9
LOCUS B1598458 817 bp mRNA linear EST 07-SEP-2001
DEFINITION 603246969F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5294681 5', mRNA sequence.

ACCESSION B1598458
VERSION B1598458.1 GI:15491397
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 817)
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM1745 row: 1 column: 18
High quality sequence stop: 386.
Location/Qualifiers
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5294681"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTV-3', size-selected for average insert size 2.3 kb and normalized to R0F 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 158 a 223 c 316 g 120 t

ORIGIN

Query Match 59.8%; Score 299.2; DB 10; Length 817;
Best Local Similarity 92.3%; Pred. No. 5.2e-41;
Matches 349; Conservative 0; Mismatches 23; Indels 6; Gaps 3;

QY 10 GCGCGGTTGTGAGTACCGGAGTGGGATGATCCCGGCTAGGGAGCGGCGC--C 67
|||||
Db 20 GCGCGGTTGTGAGTACCGGAGTGGGATGATCCCGGCTAGGGAGCGGCGC 79
|||||

QY 68 CCGATCGGCTTAGTGGAGCTCCGAAGGAGTGACTAGACACCCGGGTGGCTACTTT 127
|||||
Db 80 CCGATCGGCTTAGTGGAGCTCCGAAGGAGTGACTAGACACCCGGGTGGCTACTTT 139
|||||

QY	128	TCCTCCGCTCCTTTTGGTTTTTTTTCCTTTGGGCTCGGGCTGATGTCGCCACTGAGC	187
QY	128	TCCTCCGCTCCTTTTGGTTTTTTTTCCTTTGGGCTCGGGCTGATGTCGCCACTGAGC	187
Db	140	TCCTCCGCTCCTTTTGGTTTTTTTTCCTTTGGGCTCGGGCTGATGTCGCCACTGAGC	198
QY	188	AAAGATTCCTCTGTAAAAACCAGAGCAGCCCTCCCTCAATTGTTGGGCTCGGGAGTGTG	247
Db	199	AAAGATTCCTCTGTAAAAACCAGAGCAGCCCTCCCTCAATTGTTGGGCTCGGGAGTGTG	258
QY	248	GCGGTGCGCCGAGCGCGCGGGGCGGAGGCAAAAGGAGGAGCAGCCGCGCGGAGCGGG	307
Db	259	GCGGTGCGCCGAGCGCGCGGGGCGGAGGCAAAAGGAGGAGCAGCCGCGCGGAGCGGG	318
QY	308	CCTCGAGCTTCCCTGCGCTCTGCTGCTGCGCCAGCGGGTTCGCTCGCTGAGAGCGCA	364
Db	319	CCTCGAGCTTCCCTGCGCTCTGCTGCTGCGCCAGCGGGTTCGCTCGCTGAGAGCGC	378
QY	365	GCGGCGCGCGGATGAGG	382
Db	379	AGGGGCGCGCGGATG	396

RESULT	10
AM028893	
LOCUS	310 bp mRNA linear EST 20-OCT-2000
DEFINITION	wv97g12.x1 NCI_GGAP_Gas4 Homo sapiens CDNA clone IMAGE:2542102 3'.
ACCESSION	AM028893
VERSION	AM028893.1 GI:5867660
KEYWORDS	EST.
SOURCE	human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 310)
REFERENCE	NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap .
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 298 Std Error: 0.00
 Seq primer: -40UP from Gldco
 High quality sequence stop: 302.
 Location/Qualifiers
 1. 310
 FEATURES
 source

FEATURES	SOURCE	Location/Qualifiers
		1. .310
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:2542102"
		/clone_1ib="NCI_CGAP_Gas4"
		/tissue_type="poorly differentiated adenocarcinoma with signet ring cell features"
		/lab_host="DH10B"
		/note="Organ: Stomach; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"
BASE COUNT		39 a 102 c 110 g 58 t 1 others
ORIGIN		

Query Match	58.6%;	Score 293.2;	DB 9;	Length 310;
Best Local Similarity	98.4%;	Pred. No. 6e-40;		
Matches 306;	Conservative	0;	Mismatches	4;
			Gaps	1;
QY	99	GTCACATGACACCCGGGGGCTACTTTTCTTCCGGTGTCTTTTTCCTTTT	158	

Db	1	GTGACTAGACACCCGGGTGGGGTACATTTTCTCCGGTGGTTTG-C-TTTTATTTCCTT	59
QY	159	GGGCTGGGGCTGAGTGTGCGCCACTAGACAAATTCCTCCGTAACCAACGAGGAGCC	218
Db	60	GGGTCTGGGGTGATGTGCGCCACTAGACAAATTCCTCGTAACCAACGAGGAGCC	119
QY	219	TCCCGTAAATTTGTTGGGCTCGGGAGTGTGCGGTGCCCGAGCGCGCGGGCGGAGGC	278
Db	120	TCCCGTAAATTTGTTGGGCTCGGGAGTGTGCGGTGCCCGAGCGCGCGGGCGGAGGC	179
QY	279	AAAGGAGCGGAGCGCGGCGCGGAGCGCGGAGCGTTCGTGCTCCTCGCTGAGCC	338
Db	180	AAAGGAGCGGAGCGCGGCGCGGAGCGCGGAGCGTTCGTGCTCCTCGCTGAGCC	239
QY	339	CAGGGGTTTCCCTGCGGTAGAGCGCGAGGCGCGCGCATGAAGCGGTGAGCCCGGTGCG	398
Db	240	CAGGGGTTTCCCTGCGCGAGGCGCGAGGCGCGCGCATGAAGCGGTGAGCCCGGTGCG	299
QY	399	CGCTTCGGGGCC 409	
Db	300	CGCTTCGGGGCC 310	

RESULT 11					
BF711180/c					
LOCUS	BF711180	455 bp	mRNA	linear	EST 02-JAN-2001
DEFINITION	MI-P-AV1-nrc-h-10-0-UI s1 MI-P-AV1 Sus scrofa cDNA clone				
	MI-P-AV1-nrc-h-10-0-UI 3', mRNA sequence.				

ACCESSION	BF711180
VERSION	BF711180.1
KEYWORDS	EST.
SOURCE	EST.
ORGANISM	Pig.
REFERENCE	Sus. scrofa
AUTHORS	Eunaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
	1 (bases 1 to 455)
	Bonaldo,M.F., Lennon,G. and Soares,M.B.
	Normalization and subtraction: two approaches to facilitate gene
	discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
COMMENT	Contact: Tugale CK

COMMENT
Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildeer Hall, Ames, IA 50011-3150, USA
Tel: 5152844252
Fax: 5152942401
Email: cktugle@iastate.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares lab, University of Iowa Esr sequencing: M.B. Soares Lab,
University of Iowa clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-No.

```

FEATURES
source
Location/Qualifiers
1..455
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-A11-nrc-h-10-0-UI"
/clone_lib="MI-P-A11"
/lab_host="DH10B (Life Technologies)"
/notes="vector: pUT3b-pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: EcoRI. The MI-P-A10
library is normalized library derived from the MI-P-A10
library, ultimately derived from placenta tissue. For a
detailed description of the library from which this clone
was derived, please visit our web site at
http://pigdb.genome.jaxstate.edu/. The procedure used to
create this library has been previously described (Bonaldo
, Lennon and Soares, Genome Research 6 : 791-806, 1996)
TAG_SEQ=None found"

```


Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401

Email: cktunglee@iastate.edu
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,
University of Iowa Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-No.

FEATURES

Location/Qualifiers
1..440
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="MI-P-AV0-nev-h-05-0-UI"
/clone_1ib="MI-P-AV0"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: EcoRI; The MI-P-AV0
library is derived from placenta. For a detailed
description of the library from which this clone was
derived, please visit our web site at
http://piglet.genome.iastate.edu/. The procedure used to
create this library has been previously described (Bonaldo
, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 62 a 168 c 145 g 65 t

ORIGIN

Query Match 49.7%; Score 248.4; DB 10; Length 440;

Best Local Similarity 79.0%; Pred. No. 1.8e-32;

Matches 346; Conservative 0; Mismatches 61; Indels 31; Gaps 3;

80 AGCGAGGCTCCGAGGAGTACTAGACACCCGGGCTACTTTCTCCGGTCT 139
|||||
428 AGCGAGGCTCCGAGGAGTACTAGACACCCGGGCTACTTTCTCCGGTCT 369
|||||
140 TTTGCTTTTCTCTTGGCTGCGGCTGAGTGTCCCACTGACCAAGATTCCCTC 199
|||||
368 ATTGATTCTTCTCCAAATGATGCCG-----GCCACAGACAGATCCCTC 319
|||||
200 GTAAACCCGAGGAGCCCTCCGTCATTTGGGCTGGGAGTGTCCG----- 250
|||||
318 GCAAAAGCGAGGAGCCCTCCGTCATTTGGGCTGGGAGTGTCCG----- 259
|||||
251 -----GTGCCCGAGCGCGCGGCGGAGGCAAAAGGAGCGAGCGCGC 299
|||||
258 GTGCCCGAGCGCGTGAACCGAGCGGCGGAGGCAAAAGGAGCGAGCGCGC 199
|||||
300 -GGAGGCGCGCGAGCTTGCCTGCTCCCTGCTCCCGGAGGCTGCGCTAG 358
|||||
198 GGGTGGGCGCGGAGCTTGCCTGCTCCCTGCTCCCGGAGGCTGCGCTAG 139
|||||
359 AGCGAGGCGCGCGATGAAGCGGCTGAGCCGCTGCGCCCTCGGCGCGAAGCGC 418
|||||
138 AAGCGAGGCGCGCGATGAAGCGGCTGAGCCGCTGCGCCCTCGGCGCGTAAGCGC 79
|||||
419 CGTGGGCTGCGCGCGGAGGAGCTGCGCTGCTGCGCGGAGCGGAGCGCAAGCC 478
|||||
78 CGTGGGCTGCGCGCGGAGGAGCTGCGCGCGCTGCTGCGCGAGCGCAAGCC 19
|||||
479 TGGTGGCTCGCGCGCG 496
|||||
18 TGGCGGCTCGCGCGCG 1

RESULT 14

BI398600/c 385 bp mRNA linear EST 14-AUG-2001

LOCUS MI-P-AV1-npu-e-10-0-UI.s1 MI-P-AV1 Sus scrofa cDNA clone
DEFINITION MI-P-AV1-npu-e-10-0-UI 3', mRNA sequence.

ACCESSION BI398600 GI:15177661
VERSION BI398600.1
KEYWORDS EST.

SOURCE

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

COMMENT

Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktunglee@iastate.edu
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,
University of Iowa Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-No.

FEATURES

source

1..385
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-AV1-npu-e-10-0-UI"
/clone_1ib="MI-P-AV1"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: EcoRI; The MI-P-AV1
library is normalized library derived from the MI-P-AV0
library, ultimately derived from placenta tissue. For a
detailed description of the library from which this clone
was derived, please visit our web site at
http://piglet.genome.iastate.edu/. The procedure used to
create this library has been previously described (Bonaldo
, Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 59 a 146 c 136 g 44 t

ORIGIN

Query Match 49.1%; Score 245.4; DB 10; Length 385;

Best Local Similarity 83.0%; Pred. No. 5.8e-32;

Matches 328; Conservative 0; Mismatches 36; Indels 31; Gaps 3;

124 CTTTCTTCCGGTCTTTGCTTTTCTTTTCTTTGGGCTGGGCTGAGTGTCCGCACT 183
|||||
385 CTTTCTTCCGGTCTTTGCTTTTCTTTTCTTTGGGCTGGGCTGAGTGTCCGCACT 336
|||||
184 GAGCAAGATTCCCTGTAACCAAGAGCAACCTCCGTCATTTGTTGGGCTGGAGAG 243
|||||
335 GAGCAAGATTCCCTGTAACCAAGAGCAACCTCCGTCATTTGTTGGGCTGGAGAG 276
|||||
244 TGTGCG-----GTGCCCGAGGCGCGGCGGCGGAGGCAAGG 283
|||||
275 CGCGGAGCTTCCGCGTGGCCCGAGCGCTGAGCGCAACGCGGCGGAGGCAAGG 216
|||||
284 GAGCGAGCGCGCGCGC-GAGCGGCGCGGAGCTTGCCTGCTCCCTGCTGCGCCAGC 342
|||||
215 GAGCGAGCGCGCGCGCGGCGGCGGAGCTTGCCTGCTCCCTGCTGCGCCAGC 156
|||||
343 GGGTTCGCTCGCGTGAAGCGAGGCGCGCGGATGAAGGCGGTGAGCCGCTGCGCCC 402
|||||
155 GGGTTCGCTCGCGTGAAGCGAGGCGCGCGGATGAAGGCGGTGAGCCGCTGCGCCC 96
|||||
403 TCGGCGCGCAAGCGCGCTGCGGCTGCGGCGGCGGAGGAGTGTGGGCTGCGTGGGC 462

Isolated polypeptide, used to treat or prognose a disorder characterized by expression of a hCAP e.g. cancer, is encoded by an isolated nucleic acid comprising an NA Group 3 or 4 molecule -

Claim 57; Page 91-92; 152pp; English.

The invention relates to nucleic acids and encoded polypeptides which are cancer associated antigens expressed in patients afflicted with small cell lung cancer. The molecules provided by the invention can be used in the diagnosis, monitoring, research or treatment of conditions characterised by the expression of one or more cancer associated antigens. The polypeptide is used to treat a disorder characterised by expression of a hCAP, and determine regression, progression or onset of a condition characterised by expression of an abnormal amount of a protein encoded by a nucleic acid (NA) group 1 molecule. The disorders are small and non-small cell lung cancer, melanoma, colon, breast, head and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma. The present sequence is a DNA encoding human ID4 protein. This small cell lung cancer associated gene is designated as NF-SCLC-4. ID4 protein is a member of dominant negative helix-loop-helix (HLH) proteins.

Sequence 1322 BP; 267 A; 376 C; 417 G; 262 T; 0 other;

Query Match 100.0%; Score 500; DB 22; Length 1322;
Best Local Similarity 100.0%; Pred. No. 6, 3e-95;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGAATTCGGGCGCGGTTGAGTAGTACCGGAGTGGGTGATCCCGGAGTGGGAGC 60
    |||||||
Db 1 GGAATTCGGGCGCGGTTGAGTAGTACCGGAGTGGGTGATCCCGGAGTGGGAGC 60
QY 61 GCGGCGCGCGATCGGCTTGTAGTCCGAAAGGAGTACTAGACACCGGCTGGG 120
    |||||||
Db 61 GCGGCGCGCGATCGGCTTGTAGTCCGAAAGGAGTACTAGACACCGGCTGGG 120
QY 121 CTACTTTCTCCGCTTGTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 180
    |||||||
Db 121 CTACTTTCTCCGCTTGTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 180
QY 181 ACTGAGCAAAAGATTCCTCTGTAACCCAGAGGAGCCCTCCCTCAATTGTTGGCTCG 240
    |||||||
Db 181 ACTGAGCAAAAGATTCCTCTGTAACCCAGAGGAGCCCTCCCTCAATTGTTGGCTCG 240
QY 241 GAGTGTGCGGCTGCGCGAGCGCGCGGCGGAGGCAAGGAGCGCGCGCGCG 300
    |||||||
Db 241 GAGTGTGCGGCTGCGCGAGCGCGCGCGGCGGAGGCAAGGAGCGCGCGCGCG 300
QY 301 GACGGGCGCGGAGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
    |||||||
Db 301 GACGGGCGCGGAGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 GCGAGGCGCGCGGATGAAGGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCG 420
    |||||||
Db 361 GCGAGGCGCGCGGATGAAGGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 TCGGGCTGCGGCGCGGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
    |||||||
Db 421 TCGGGCTGCGGCGCGGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGTGGCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500
    |||||||
Db 481 GGTGGCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500

```

RESULT 2
AAC10488
ID AAC10488 standard; cDNA; 225 BP.
XX
AC AAC10488;
XX
DT 06-Oct-2000 (first entry)
XX

Human secreted protein 5' EST, SEQ ID NO: 14563.
Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
gene therapy; chromosome mapping; ss.

Homo sapiens.
EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI; 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -

Claim 1; SEQ ID 14563; 71bp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or poly(A) RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

Sequence 225 BP; 36 A; 55 C; 80 G; 53 T; 1 other;

Query Match 36.9%; Score 184.4; DB 21; Length 225;
Best Local Similarity 98.1%; Pred. No. 1, 4e-29;
Matches 208; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

```

QY 10 GCGCGGTTGAGTAGTACCGGAGTGGGTGATCCCGGCTAGGAGCGCGCGG-C 67
    |||||||
Db 15 GCGCGGTTGAGTAGTACCGGAGTGGGTGATCCCGGCTAGGAGCGCGCGCGG 74
QY 68 CCGATGCGGCTTACCTGCGAGCTCGAAGGAGTACTAGACACCGGCTGCTCTT 127
    |||||||
Db 75 GCGATGCGGCTTACCTGCGAGCTCGAAGGAGTACTAGACACCGGCTGCTCTT 133
QY 128 TCTTCGCGCTTGTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 187
    |||||||
Db 134 TCTTCGCGCTTGTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 193
QY 188 AAGATTCCTCTGTAACCCAGAGGAGCGACCT 219
    |||||||
Db 194 AAGATTCCTCTGTAACCCAGAGGAGCGACCT 225

```

RESULT 3
AA094779
ID AA094779 standard; DNA; 1309 BP.
XX
AC AA094779;
XX
DT 07-MAR-1996 (first entry)
XX

DE Regulator of neuronal proliferation RNP-1 partial DNA.
 XX
 KM Regulator of neuronal proliferation; neuron; RNP-1; ss.
 XX Homo sapiens.
 OS
 XX MO9516774-A1.
 PN
 XX 22-JUN-1995.
 PD
 XX 19-DEC-1994; 94WO-US14614.
 PF
 XX 08-SEP-1994; 94US-031416.
 PR 17-DEC-1993; 93US-0169522.
 PA (SPIN-) SPINAL CORD SOC.
 XX Neuman T, Nornes HO, Suda K;
 XX WPI: 1995-231568/30.
 DR
 XX A vector contg. a nucleic acid encoding an E2F and/or E1A regulator
 PP - is used to induce DNA synthesis in differentiated neurons and
 PP replace neurons after injury or neurodegenerative diseases.
 PS
 XX Disclosure; Page 30; 60pp; English.
 CC This sequence corresponds to RNP-1 which has no significant
 CC homology to other characterized sequences.
 XX Sequence 1309 BP; 305 A; 339 C; 381 G; 284 T; 0 other:
 XX

```

Query Match          23.4%; Score 116.8; DB 16; Length 1309;
Best Local Similarity 61.88; Pred no. 1.4e-15;
Matches 220; Conservative 0; Mismatches 132; Indels 4; Gaps 2
Oy      145 TTTTCTTCCTTGCGGCGGTGAGATGTCGGCCACTGACCAAAAGATTTCCTCCTGA
Db      1 TTTTCTTCCTTGCTTGCGGCGGTGAGATGTCGGCCACTGACCAAAAGAAAAAACAAA
Oy      205 ACCGAGAGCACCCGCCGCATTTGTGGGTGGGAGTGGCGGCGCCCGAGCTTCGCTGC
Db      61 ACATAAACCTGACACTACCCCTTCCAAAACCCTGACAGCTCTGCTGCAATTTCG--GAG
Oy      265 CGGGGCGCGGAGGCAAAAGAGCGGAGCGCGCGCGGAGCGGGGCCCCGGAGCTTCGCTGC
Db      118 CTCGAGCCCGCGCGGCTGCTGSAAGCCGCGAGACGAGAGAGGAGCAAAGCAAAATTCG
Oy      325 TCCTCTCGGTGCGCCCGACGCGGTTTCCGCTGACGACAGCGCGAGCGCGCGATGAAGCG
Db      178 GCGGGGAGCGACCTGCGCTGCTGTCGTCGCTGCTCCAGAACGCGCGATGAAGCGG
Oy      385 GTGACACCGGTGTGCGCCCTCGGAGCCGACAAGAGCGCGCTGCGGCGGTGCGGCGGAGCTG
Db      238 GTGACACCGCGGTGCG--CCCTCGGAGTCAAGAGCGCGCTGCGGCGGTGCGGCGGAGCTG
Oy      445 GCGCTGCGCTGCGCTGAGCGACAGCGCAACACTGGTGGCTCGCGACGCCGCGGCG 500
Db      297 GCGGTACCGCTGCGCGCGAGCAAGCGCCACAGCCTGGGTGGCTGCGGACGCCGCGC 352

```

RESULT 4
 AAD25519 ID AAD25519 standard; DNA; 154746 BP.
 XX
 XX AAD25519;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX Human herpesvirus 2 complete DNA genome.
 DE
 XX Human herpesvirus 2; cytosolic; cancer; immunosuppressive; virucide;
 KM antibacterial; fungicide; protozoacide; antitubercular; antiinflammatory;

KW antiaerobic; rheumatoid arthritis; neuroprotective; multiple sclerosis;
 KW immune response; vasotropic; vaccine; gene therapy; autoimmune disease;
 KW vasculitis; ds.
 OS Human herpesvirus 2.
 XX
 XX MO200176643-A1.
 PN 18-OCT-2001.
 PD 06-APR-2001: 2001MO-US11372.
 PE 07-APR-2000: 2000US-195680P.
 PR (BAYU) BAYLOR COLLEGE MEDICINE.
 PA Orson FM, Kinsey BM, Bhogal BS;
 PI WPI: 2002-066308/09.
 DR
 XX Composition for oral delivery of vaccines, comprises expression vector
 PT containing antigenic genomic sequence, bound to aggregated
 PT protein-polycationic polymer conjugate or suspension
 XX
 XX disclosure: Page 90-132: 145pp: English.
 XX
 CC The invention relates to a composition comprising an expression vector
 CC bound to an aggregated protein-polycationic polymer conjugate or
 CC suspension. The expression vector contains a promoter polynucleotide
 CC sequence operatively linked to a polynucleotide sequence encoding an
 CC antigen which is a fragment of a gene or genome associated with an
 CC infectious disease and the antigenic disease associated with an
 CC infectious disease is a bacterial disease, a viral disease, a fungal
 CC disease, a parasitic disease, a protozoan disease, a helminth disease,
 CC a disease of bacterium, a fungus, a protozoa and virus such as human
 CC immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C
 CC virus (HCV), influenza and respiratory syncytial virus (RSV), and
 CC optionally comprising a nucleotide sequence encoding a cytokine (or a
 CC cytokine expression vector), is useful for inducing an immune response
 CC (systemic and/or mucosal) in an organism. The cytokine expression vector
 CC contains a sequence for granulocyte macrophage colony stimulating factor
 CC (GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding
 CC the antigen and the cytokine are under transcriptional control of same or
 CC different promoter polynucleotide sequences. The expression vector, as a
 CC DNA vaccine is useful for treating a condition in an organism. The
 CC present sequence is human herpesvirus 2 complete DNA genome related
 CC to the invention.
 CC
 XX
 XX Sequence 154746 BP: 23003 A: 54218 C: 54701 G: 22824 T: 0 other:
 SO
 Query Match 10.3% Score 51.6: DB 24: Length 154746:
 Best Local Similarity 50.3% Pred. No. 0.051:
 Matches 155: Conservative 0: Mismatches 149: Indels 4: Gaps 1
 OY 191 GATTCCTCCCTAAAAACCCAGAGACACCTCCCGTCAATTTTGAGGCTGTGGAGTGTGTGCG 250
 DB 1124 GGTGGCTCTGGAGAACGACGCGCGCGCCCTGTGCTCCGACGCGGGTCTCTGGGCGCCCAAGGCGGC 1183
 OY 251 GGTGCCCCGAGCGCGCCGCGCGAGGCAAAAGGAGGAGCGAGCGCCGCGAGCGGAGCC 310
 DB 1184 CGAGCGACACGGGTTCGCGGCGCCCGGTGCGGCGCGAGGCGGTCATCGAGGACGTCT 1243
 OY 311 GGAGCTTGCCTGCTCTCCCTGCTGCTGTGCCGCCAGCGGGTTCGCTGACGCGAGGCGCGC 370
 DB 1244 GGAGGCTCCAGAGCCCGAGGCTCGAGGCCCGAGGCCGAGCCCGGCGCCACGAAAGACGCGCGAC 1303
 OY 371 GGCGGANTGAAGCGGATGAGCCGCTGCGGCTCTCGGGGCGCCAAAGCGCGCTGTGGGCTGCG 430
 DB 1304 CGCGGAGAGAGAGAGAGCGCGCGCGGCGGCGGAGGCTCTCGCGCGCGGCGGAGCCCGG 1363
 OY 431 GGCGCGGAGAGACTGTGCGCTGCTGCTGCGCGAGACGCCACAGCCTGTGGGTGTGCTCG 490
 DB 1364 CGGTGGG---CGGTGAGGAGGTTGAACCGGAGAGGAGGCGGCTGTGCGCGGAGAGCCCG 1419

OY	491	CACCGCG	498
Db	1420	GAGCTCG	1427
	RESULT 5		
ID	AAD25519/c		
XX	AAD25519 standard; DNA: 154746 BP.		
XX			
XX	AAD25519;		
DT	26-MAR-2002 (first entry)		
DE	Human herpesvirus 2 complete DNA genome.		
XX			
KW	Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;		
KW	antibacterial; fungicide; protozoacide; antithrombotic; antiinflammatory;		
KW	atherosclerotic; rheumatoid arthritis; neuroprotective; multiple sclerosis;		
KW	immune response; vasotropic; vaccine; gene therapy; autoimmune disease;		
XX	vaccillitis; ds.		
OS			
XX	Human herpesvirus 2.		
XX	MO2001I76643-A1.		
XX			
XX	18-OCT-2001.		
XX			
XX	06-APR-2001; 2001MO-US11372.		
XX			
PR	07-APR-2000; 2000US-195680P.		
PA			
PA	(BAYU) BAYLOR COLLEGE MEDICINE.		
EI	Orson FM, Kinsey BM, Bhogal BS;		
DR	WPI: 2002-066308/09.		
XX			
PT	Composition for oral delivery of vaccines, comprises expression vector		
PT	containing antigenic genomic sequence bound to aggregated		
PT	protein-polycationic polymer conjugate or suspension		
PS	Disclosure; Page 90-132; 145pp; English.		
CC	The invention relates to a composition comprising an expression vector		
CC	bound to an aggregated protein-polycationic polymer conjugate or		
CC	suspension. The expression vector contains a promoter polynucleotide		
CC	sequence operatively linked to a polynucleotide sequence encoding an		
CC	antigen which is a fragment of a gene or genome associated with an		
CC	infectious disease, cancer, and autoimmune disease such as rheumatoid		
CC	arthritis, vasculitis, and multiple sclerosis, pathogenic genomes		
CC	consisting of bacterium, fungus, protozoa and virus such as human		
CC	immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C		
CC	virus (HCV), influenza and respiratory syncytial virus (RSV), and		
CC	optionally comprising a nucleotide sequence encoding a cytokine (or a		
CC	cytokine expression vector), is useful for inducing an immune response		
CC	(systemic and/or mucosal) in an organism. The cytokine expression vector		
CC	contains a sequence for granulocyte macrophage-colony stimulating factor		
CC	(GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding		
CC	the antigen and the cytokine are under transcriptional control of same or		
CC	different promoter polynucleotide sequences. The expression vector, as a		
CC	DNA vaccine is useful for treating a condition in an organism. The		
CC	present sequence is human herpesvirus 2 complete DNA genome related		
CC	to the invention.		
XX			
XO	Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 other;		
Query Match	10.3%; Score 51.6; DB 24: Length 154746;		
Best Local Similarity	50.3%; Pred.No. 0.051;		
Matches 155; Conservative	0; Mismatches 149; Indels 4; Gaps 1.		
OY	191	GATTCCTGCTAAACCCAGAGCACCTTCCGTCATTTGTGGCTCCTGGACTGTCGCG	250
Db	126126	GGTAGCGTGGAAGACGCCGCAGCCCTGCAAGAGGAGGAGGCTCTGAGCGGCGCACAGGAGAC	126067

[illegible]

PT diagnosing and/or identifying therapeutic agents for atherosclerosis
 XX
 PS Claim 4; Fig 2A; 143pp; English.
 XX
 CC The present sequence is that of cDNA encoding novel rabbit
 CC low density lipoprotein binding protein 1 (LBP-2, see AAB2807).
 CC The cDNA was isolated following screening of a rabbit cDNA library
 CC for clones encoding LBPs that bound to both native low density
 CC lipoprotein (LDL) and methyl LDL. The invention provides claimed
 CC polynucleotides encoding novel polypeptides which are capable of
 CC binding to native and methylated LDL, the isolated polypeptides,
 CC termed LBPs, and biologically active fragments and analogues of
 CC them, as well as expression vectors, cells and methods of producing
 CC the LBPs. Also claimed are methods for evaluating an agent for
 CC at risk for atherosclerosis, methods for evaluating an agent for
 CC use in treating atherosclerosis, and methods for treating a cell
 CC having an abnormality in structure or metabolism of LBP.
 CC Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed.
 CC
 XX Sequence 2561 BP; 372 A; 937 C; 879 G; 373 T; 0 other;
 SQ
 Query Match 9.8%; Score 49; DB 22; Length 2561;
 Best Local Similarity 54.0%; Pred. No. 0.16;
 Matches 143; Conservative 0; Mismatches 120; Indels 2; Gaps 2;
 QY 233 GGGCTGGGAGTGTGCGGTGCTCCCGAGCGCGCGCGGAGCAAGGAGCGGAGC 292
 DB 863 GGGCGGG 804
 QY 293 CGGGCGGGGAGGG 352
 DB 803 GGG 744
 QY 353 GCGTAGAGCGGAGGCGGCGCGATGAGGCGGTGAGCGCGGTGCGCGCGCGCGCA 412
 DB 743 G-GGGGGCGGGAGGGGGGGGGCGCTGTGGCGGAGCGGCGGGCGG-GCGGCGGGGGCGCC 686
 QY 413 AGCGCGCGTGGCGGTGCGGGCGGGGGAGCTGCGCGCTGCTGGCGGAGCGAGCGGC 472
 DB 685 CGGGCGGG 626
 QY 473 ACAGCGTGGGTGCGTCCGACGCGC 497
 DB 625 GCGGGCGCGGG 601
 RESULT 9
 AAX53491/C
 ID AAX53491 standard; DNA; 114955 BP.
 XX
 AC AAX53491;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE Human adenosine A1 receptor antisense oligonucleotide fragment.
 XX
 XX Antisense oligonucleotide; multiple target; antisense treatment;
 XX impaired respiration; inflammation; lung disease;
 XX pulmonary vasoconstriction; inflammation; allergic rhinitis;
 XX acute asthma; allergy; asthma; impeded respiration;
 XX respiratory distress syndrome; pain; cystic fibrosis;
 XX pulmonary hypertension; pulmonary vasoconstriction; emphysema;
 XX chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
 XX colon cancer; breast cancer; lung cancer; pancreatic cancer;
 XX hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
 XX prostate cancer; ss.
 XX
 OS Synthetic.
 XX
 AC WO9113886-A1.
 PN
 DT 25-MAR-1999.
 PD

XX
 PF 17-SEP-1998; 98WO-US19419.
 XX
 PR 09-JUN-1998; 98US-0093972.
 PR 17-SEP-1997; 97US-0059160.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 XX
 PI Nyce JW;
 XX
 DR WPI; 1999-229400/19.
 XX
 PT New antisense oligonucleotides used in treatment of, e.g. pulmonary
 PT vasoconstriction
 PS
 PS Disclosure; Page 37; 120pp; English.
 XX
 CC The specification describes antisense oligonucleotides (AAX52869-X55271)
 CC directed against at least 2 mRNAs selected from target genes, coding and
 CC non-coding regions of RNAs corresponding to target genes, gene
 CC initiation codons, genomic flanking regions, intron-exon borders, the
 CC 5'-end, the 3'-end and the junction between coding and non-coding
 CC regions and all segments of RNAs encoding proteins associated with one
 CC or more diseases, conditions or mixtures. The antisense oligonucleotides
 CC may be derived from sequences AAX55272-74. These multiple target
 CC oligonucleotides (specifically AAX55180-271) can be used for the
 CC antisense treatment of diseases and conditions. Typical diseases and
 CC conditions are those associated with impaired respiration and
 CC inflammation, including lung diseases, pulmonary vasoconstriction,
 CC inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded
 CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
 CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
 CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
 CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
 CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
 CC hepatic metastases, as well as all types of cancers which may metastasize
 CC or have metastasized to the lungs, including breast and prostate cancer.
 CC
 XX Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;
 SQ
 Query Match 9.8%; Score 48.8; DB 20; Length 114955;
 Best Local Similarity 37.4%; Pred. No. 0.19;
 Matches 111; Conservative 22; Mismatches 162; Indels 2; Gaps 1;
 QY 206 CCCAGAGCACCCTCCGCAATTGTGGGTGCGGAGTGTGCGGTGCCCCGAGCGGC 265
 DB 110735 CCCGGGCCCGGCGCGCCCGCCGNNNNNNNSCCGGCGGCCCGCCGNNNNNSGG 110676
 QY 266 CGGGCGGAGGCAAGGAGCGGAGCGCGCGCGGAGCGGGCCCGGAGCTTGCCT 325
 DB 110675 CGGGCGGNNNNNSGGCGGGCGGNNNNNSGGCGGGCGGCGNNNNNSGGCGGGCG 110616
 QY 326 CCCTGCTGCGCCCGACGCGGGTTCGCTGCGTAGAGCGAGCGCGCGCGATGAGGCGG 385
 DB 110615 GCGCGNNNNNSGGCGGGCGGCGGNNNNNSGGCGGGCGGCGGNNNNNSGGCGGG 110556
 QY 386 TGAGCCCGGTGCGCCCT--CGGGCGCAAGCGCGCTGCGGTGCGGCGGGAGCT 443
 DB 110555 GCGGGCGGNNNNNSGGCGGGCGGCGGCGGCGGNNNNNSGGCGGGCGGGCGCC 110496
 QY 444 GCGGCTGCGTGTGCTGCGGCGAGCAGCGCCACAGCGCGGTGCGCGACGCGCGC 500
 DB 110495 NNNNNNSGGCGGGCGGCGGCGCCNNNNNSGCGCGGGGCGCCNNNNNSGGCGCG 110439
 RESULT 10
 AAH81775/C
 ID AAH81775 standard; DNA; 3787 BP.
 XX
 AC AAH81775;
 XX
 DT 21-SEP-2001 (first entry)
 XX


```
FT /standard_name= "Single nucleotide polymorphism"
FT replace(33840,A)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT 33881..35300
FT /tag= u
FT /number= 6
FT /cons_splice=(5'site:NO,3'site:NO)
FT replace(33922,G)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT replace(35202,G)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT 35301..35343
FT /tag= x
FT /number= 7
FT 35344..37315
FT /tag= y
FT /number= 7
FT replace(35497,G)
FT /tag= z
FT /standard_name= "Single nucleotide polymorphism"
FT replace(37287,A)
FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT replace(37296,A)
FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT 37316..37427
FT /tag= ac
FT /number= 8
FT 37428..38381
FT /tag= ad
FT /number= 8
FT /cons_splice=(5'site:YES,3'site:NO)
FT replace(37545,A)
FT /tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT 38382..38517
FT /tag= af
FT /number= 9
FT replace(38397,T)
FT /tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT 38518..40495
FT /tag= ah
FT /number= 9
FT replace(38622,C)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT 40496..40562
FT /tag= aj
FT /number= 10
FT replace(40516,C)
FT /tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT replace(40521,T)
FT /tag= al
FT /standard_name= "Single nucleotide polymorphism"
FT 40563..40748
FT /tag= am
FT /number= 10
FT 40749..40861
FT /tag= an
FT /number= 11
FT replace(40841,A)
FT /tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT replace(40875,C)
FT /tag= ap
FT /standard_name= "Single nucleotide polymorphism"
XX
```

```
PN WO200190125-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US16909.
XX
XX 24-MAY-2000; 2000US-206664P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Chew A, Duda A, Koshy B;
XX
XX WPI: 2002-106169/14.
XX P-PSDB; AAU11951.
XX
XX Isolated human pyridoxal (pyridoxine, vitamin B6) kinase polyNTs,
XX useful for therapeutic purposes, for studying the expression and
XX function of the polyNT, and for expressing pyridoxal protein -
XX
XX Disclosure: Fig 1; 135pp; English.
```

The invention describes an isolated human pyridoxal (pyridoxine, vitamin B6) kinase, (PDXK) polynucleotide. The polynucleotide is useful in studying the expression and function of PDXK, and in expressing PDXK protein for use in screening for candidate drugs to treat PDXK related diseases and for therapeutic purposes. A transgenic animal is useful for studying expression of the PDXK isogenes in vivo, for in vivo screening and testing of drugs targeted against PDXK protein, and for testing the efficacy of therapeutic agents and compounds for autoimmune polyglanular disease type 1. The polypeptide is useful for studying the effect of the variation on the biological activity of PDXK and the binding affinity of candidate drugs targeting PDXK for the treatment of autoimmune polyglanular disease type 1. Genotyping and haplotyping is useful for improving the efficacy and reliability of several steps in the discovery and development of drugs for treating diseases associated with PDXK activity, e.g., autoimmune polyglanular disease type 1, to validate PDXK as a candidate agent for treating a specific condition or disease CC predicted to be associated with PDXK activity, and in the design of CC clinical trials of candidate drugs. This sequence (located on chromosome CC 21q22.3) encodes human pyridoxal (pyridoxine, vitamin B6) kinase (PDXK) CC and forms the reference sequence on which the isoforms AAU11952 and CC AAU11953 are based, described in the method of the invention.

Sequence 44861 BP; 9041 A; 12297 C; 13398 G; 10125 T; 0 other;

Query Match 9.4%; Score 47.2; DB 24; Length 44861;

Best Local Similarity 50.0%; Pred. No. 0.4; Mismatches 118; Indels 0; Gaps 0;

Matches 118; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```
QY 236 CTCGGAGTGTGGCGTGCCTCCGAGCGCCGCGCGCGAGCAAGAGGAGCGCGG 295
   ||||| ||||| | | | | | ||||| || | | | | |
DB 26041 CCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCG 26100

QY 296 CCGCGAGCGGGCCCGGAGCTTGCCTGCTCCTCGCTGCGCCAGCGGGTTCGTCGCG 355
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 26101 CGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGCG 26160

QY 356 TTAGCGGAGCGCGCGCGCGGATGAAGCGGTAGACCCCGTGCCTCGCGCGCGCAAG 415
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 26161 GGGGGCTGGCGGGGGGCTGGCGGGGGGCTGGCGGGGGGCTGGCGGGGGGCTGGCGGGG 26220

QY 416 GCGCGTGGGCTGCGCGGCGGAGCTGGCGTGCCTGCGCGGAGCAAGCGG 471
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 26221 GGGCTCGCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGCGGGGGGCTCGAGGC 26276
```

RESULT 12
AAV62176/c
ID AAV62176 standard; DNA; 117213 BP.
AC AAV62176;
XX
XX 13-JAN-1999 (first entry)
DT

```
XX HSV-2 strain SB5 Contig ID 15 DNA sequence.
XX
XX HSV-2 strain SB5: immunological response induction; therapy;
XX antiviral identification; viral protein inhibitor; ss.
XX
XX Herpes simplex virus type 2.
XX
FH Key Location/Qualifiers
FT CDS 755..1297
FT     /tag= a
FT     /product= "ORF#1 protein"
FT     /note= "encoded protein shown in AAW72170"
FT     1170..2174
FT     /tag= b
FT     /product= "ORF#2 protein"
FT     /note= "encoded protein shown in AAW72171"
FT     2229..2930
FT     /tag= c
FT     /product= "ORF#3 protein"
FT     /note= "encoded protein shown in AAW72172"
FT     complement (3130..3735)
FT     /tag= d
FT     /product= "ORF#4 protein"
FT     /note= "encoded protein shown in AAW72173"
FT     complement (3802..6447)
FT     /tag= e
FT     /product= "ORF#5 protein"
FT     /note= "encoded protein shown in AAW72174"
FT     6017..8482
FT     /tag= f
FT     /product= "ORF#6g protein"
FT     /note= "encoded protein shown in AAW72250"
FT     6026..8482
FT     /tag= g
FT     /product= "ORF#6f protein"
FT     /note= "encoded protein shown in AAW72249"
FT     6065..8482
FT     /tag= h
FT     /product= "ORF#6e protein"
FT     /note= "encoded protein shown in AAW72248"
FT     6167..8482
FT     /tag= i
FT     /product= "ORF#6d protein"
FT     /note= "encoded protein shown in AAW72247"
FT     6296..8482
FT     /tag= j
FT     /product= "ORF#6c protein"
FT     /note= "encoded protein shown in AAW72246"
FT     6326..8482
FT     /tag= k
FT     /product= "ORF#6b protein"
FT     /note= "encoded protein shown in AAW72245"
FT     6446..8482
FT     /tag= l
FT     /product= "ORF#6a protein"
FT     /transl_except= (pos: 7400..7402, aa: Ala-Ala)
FT     /transl_except= (pos: 7481..7486, aa: Ile)
FT     /note= "encoded protein shown in AAW72244"
FT     8457..9347
FT     /tag= m
FT     /product= "ORF#7 protein"
FT     /note= "encoded protein shown in AAW72175"
FT     complement (9604..11855)
FT     /tag= n
FT     /transl_except= (pos: 11635..11636, aa: Ala)
FT     /product= "ORF#8 protein"
FT     /note= "encoded protein shown in AAW72176"
FT     complement (11905..14508)
FT     /tag= o
FT     /product= "ORF#9b protein"
FT     /note= "encoded protein shown in AAW72222"
FT     complement (11905..14520)
FT CDS
```

```
FT     /tag= p
FT     /product= "ORF#9a protein"
FT     /note= "encoded protein shown in AAW72223"
FT     14399..15802
FT     /tag= q
FT     /product= "ORF#10 protein"
FT     /note= "encoded protein shown in AAW72177"
FT     complement (15996..16286)
FT     /tag= r
FT     /product= "ORF#11 protein"
FT     /note= "encoded protein shown in AAW72178"
FT     complement (16202..18064)
FT     /tag= s
FT     /product= "ORF#12 protein"
FT     /note= "encoded protein shown in AAW72179"
FT     complement (18105..19661)
FT     /tag= t
FT     /product= "ORF#13 protein"
FT     /note= "encoded protein shown in AAW72180"
FT     complement (19415..20074)
FT     /tag= u
FT     /product= "ORF#14 protein"
FT     /note= "encoded protein shown in AAW72181"
FT     20155..21453
FT     /tag= v
FT     /product= "ORF#15 protein"
FT     /note= "encoded protein shown in AAW72182"
FT     complement (21326..22291)
FT     /tag= w
FT     /product= "ORF#16 protein"
FT     /note= "encoded protein shown in AAW72183"
FT     complement (22546..24654)
FT     /tag= x
FT     /product= "ORF#17 protein"
FT     /note= "encoded protein shown in AAW72184"
FT     24684..25955
FT     /tag= y
FT     /product= "ORF#18 protein"
FT     /note= "encoded protein shown in AAW72185"
FT     complement (26295..27251)
FT     /tag= z
FT     /product= "ORF#19 protein"
FT     /note= "encoded protein shown in AAW72186"
FT     complement (27630..31754)
FT     /tag= aa
FT     /product= "ORF#20b protein"
FT     /note= "encoded protein shown in AAW72225"
FT     complement (27630..31784)
FT     /tag= ab
FT     /product= "ORF#20a protein"
FT     /note= "encoded protein shown in AAW72224"
FT     complement (32067..32735)
FT     /tag= ac
FT     /product= "ORF#21 protein"
FT     /note= "encoded protein shown in AAW72187"
FT     33140..34984
FT     /tag= ad
FT     /product= "ORF#22a protein"
FT     /note= "encoded protein shown in AAW72226"
FT     33386..34984
FT     /tag= ae
FT     /product= "ORF#22b protein"
FT     /note= "encoded protein shown in AAW72227"
FT     complement (35205..37721)
FT     /tag= af
FT     /product= "ORF#23 protein"
FT     /note= "encoded protein shown in AAW72188"
FT     complement (38058..39188)
FT     /tag= ag
FT     /product= "ORF#24 protein"
FT     /note= "encoded protein shown in AAW72189"
FT     39090..39935
FT     /tag= ah
```

```
FT /product= "ORF#25 protein"
FT /note= "encoded protein shown in AAW72190"
FT CDS
FT /tag= a1
FT /product= "ORF#26 protein"
FT /note= "encoded protein shown in AAW72191"
FT CDS
FT /tag= a1
FT /product= "ORF#27 protein"
FT /note= "encoded protein shown in AAW72192"
FT CDS
FT /tag= ak
FT /product= "ORF#28 protein"
FT /note= "encoded protein shown in AAW72193"
FT CDS
FT /tag= a1
FT /product= "ORF#29 protein"
FT /note= "encoded protein shown in AAW72194"
FT CDS
FT /tag= am
FT /product= "ORF#30 protein"
FT /note= "encoded protein shown in AAW72195"
FT CDS
FT /tag= an
FT /product= "ORF#31 protein"
FT /note= "encoded protein shown in AAW72196"
FT CDS
FT /tag= ao
FT /product= "ORF#32 protein"
FT /note= "encoded protein shown in AAW72197"
FT CDS
FT /tag= ap
FT /product= "ORF#33 protein"
FT /note= "encoded protein shown in AAW72198"
FT CDS
FT /tag= aq
FT /product= "ORF#34 protein"
FT /note= "encoded protein shown in AAW72199"
FT CDS
FT /tag= ar
FT /product= "ORF#35 protein"
FT /note= "encoded protein shown in AAW72200"
FT CDS
FT /tag= as
FT /product= "ORF#36 protein"
FT /note= "encoded protein shown in AAW72201"
FT CDS
FT /tag= at
```

Query Match 9.4%; Score 46.8; DB 19; Length 117213;
Best Local Similarity 54.2%; Pred. No. 0.5;
Matches 116; Conservative 0; Mismatches 97; Indels 1; Gaps 1;

```
QY 229 TGTTCGCTTCGAGAGTCTCGCGGCGCCGAGGCAAGGAGACG 288
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109471 TCCCGGCGGGGGGGGGGGGGGGGGGGGAAAGAAAGGCCCG 109412

QY 289 GAGCCGCGCGGAGCGGGCGGAGCTTCCTCCCTCGCCCGAGCGGGTTC 348
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109411 GAAGC-GCGGAGAGGGGCGCGGCGGCGGCGGGGGCGCGCGGC 109353

QY 349 GCTCGGTAGAGCGGCGCGCGCATGAAGCGGTGAAGCCCGCCCTCGGGC 408
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109352 CGGGCGGGGGCGCGCGCGCGGGCGCGCGCGGGCGCGCGGC 109293

QY 409 CGCAAGCGCGCTTCGCGGCTCGCGCGCGGAGC 442
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109292 CGGGCGGGGGCGCGCGCGCGGGCGCGCGGC 109259
```

RESULT 13
AAF30757/c

```
ID AAF30757 standard; DNA; 47981 BP.
XX
AC AAF30757;
XX
XX 21-JUN-2001 (first entry)
XX
XX Micromonospora megalomicea megalomicin biosynthetic gene cluster.
XX DE motillide; antiparasitic; ds.
XX
XX Megalomicin; meg gene; polyketide synthase; antibiotic;
XX
XX Micromonospora megalomicea subsp. nigra.
OS
FH Key
FH CDS Location/Qualifiers
FT complement (1..144)
FT /tag= a
FT /partial
FT /gene= "megBVI(meg1)"
FT /product= "rDP-4-keto-6-deoxyglucose-2,3-dehydratase"
FT /note= "encodes AAB82201"
FT CDS
FT /tag= b
FT /gene= "megDVI"
FT /product= "rDP-4-keto-6-deoxyhexose 3,4-isomerase"
FT /note= "encodes AAB82202"
FT CDS
FT /tag= c
FT /gene= "megDI"
FT /product= "rDP-megosamine glycosyltransferase"
FT /note= "eryCIII homologue; encodes AAB802203"
FT CDS
FT /tag= d
FT /gene= "megI"
FT /product= "mycarose O-acyltransferase"
FT /note= "encodes AAB82204"
FT CDS
FT /tag= e
FT /gene= "megDII"
FT /product= "rDP-3-keto-6-deoxyhexose 3-aminotransaminase"
FT /note= "eryCI, DnrI homologue; encodes AAB82205"
FT CDS
FT /tag= f
FT /gene= "megDIII"
FT /product= "daunosaminyl-N,N-dimethyltransferase"
FT /note= "eryCVI homologue; encodes AAB82206"
FT CDS
FT /tag= g
FT /gene= "megDIV"
FT /product= "rDP-4-keto-6-deoxyhexose 3,5-epimerase"
FT /note= "eryBVIII, dnmU homologue; encodes AAB82207"
FT CDS
FT /tag= h
FT /gene= "megDV"
FT /product= "rDP-4-keto-6-deoxyhexose 4-ketoreductase"
FT /note= "eryYIV, dnmY homologue; encodes AAB82208"
FT CDS
FT /tag= i
FT /gene= "megBII-1(megDVI1)"
FT /product= "rDP-4-keto-6-deoxyhexose 2,3-reductase"
FT /note= "encodes AAB82209"
FT CDS
FT /tag= j
FT /gene= "megBV"
FT /product= "rDP-mycarose glycosyltransferase"
FT /note= "encodes AAB82210"
FT CDS
FT /tag= k
FT /gene= "megBIV"
FT /product= "rDP-4-keto-6-deoxyhexose 4-ketoreductase"
FT /note= "encodes AAB82211"
FT CDS
FT /tag= l
```

```

FT /gene= "mega"
FT /tag= 1
FT /product= "megalomicin 6-deoxyerythronolide B
FT /note= "polyketide synthase, encodes AAB82212"
FT 12505..13470
FT /tag= m
FT /gene= "mega"
FT /function= "AT-L"
FT 13576..13791
FT /tag= n
FT /gene= "mega"
FT /function= "ACP-L"
FT 13849..15126
FT /tag= o
FT /gene= "mega"
FT /function= "KS1"
FT 15427..16476
FT /tag= p
FT /gene= "mega"
FT /function= "Ari"
FT 17155..17694
FT /tag= q
FT /gene= "mega"
FT /function= "KR1"
FT 17947..18207
FT /tag= r
FT /gene= "mega"
FT /function= "ACP1"
FT 18268..19548
FT /tag= s
FT /gene= "mega"
FT /function= "KS2"
FT 19876..20910
FT /tag= t
FT /gene= "mega"
FT /function= "AT2"
FT 21517..22053
FT /tag= u
FT /gene= "mega"
FT /function= "KR2"
FT 22318..22575
FT /tag= v
FT /gene= "mega"
FT /function= "ACP2"
FT 22867..33555
FT /tag= w
FT /gene= "mega"
FT /product= "megalomicin 6-deoxyerythronolide B
FT /note= "polyketide synthase, encodes AAB82213"
FT 22957..24237
FT /tag= x
FT /gene= "mega"
FT /function= "KS3"
FT 24544..25581
FT /tag= y
FT /gene= "mega"
FT /function= "AT3"
FT 26230..26733
FT /tag= z
FT /gene= "mega"
FT /function= "KR3 (inactive)"
FT 26998..27258
FT /tag= aa
FT /gene= "mega"
FT /function= "ACP3"
FT 27393..28590
FT /tag= ab
FT /gene= "mega"
FT /function= "KS4"
FT 28897..29931
FT /tag= ac

```

```

FT /gene= "mega"
FT /function= "AT4"
FT 29953..30477
FT /tag= ad
FT /gene= "mega"
FT /function= "DH4"
FT 31396..32244
FT /tag= ae
FT /gene= "mega"
FT /function= "ER4"
FT 32257..32799
FT /tag= af
FT /gene= "mega"
FT /function= "KR4"
FT 33052..33312
FT /tag= ag
FT /gene= "mega"
FT /function= "ACP4"
FT 33666..43271
FT /tag= ah
FT /gene= "mega"
FT /product= "megalomicin 6-deoxyerythronolide B
FT /note= "polyketide synthase, encodes AAB82214"
FT 22957..24237
FT /tag= ai
FT /gene= "mega"
FT /function= "KS5"
FT 24544..25581
FT /tag= aj
FT /gene= "mega"
FT /function= "AT5"
FT 26230..26733
FT /tag= ak
FT /gene= "mega"
FT /function= "KR5"
FT 26998..27258
FT /tag= al
FT /gene= "mega"
FT /function= "ACP5"
FT 27393..28590
FT /tag= am
FT /gene= "mega"
FT /function= "KS6"
FT 28897..29931
FT /tag= an
FT /gene= "mega"
FT /function= "AT6"
FT 29953..30477
FT /tag= ao
FT /gene= "mega"
FT /function= "KR6"
FT 31396..32244
FT /tag= ap

Query Match      9.1%; Score 45.6; DB 22; Length 47981;
Best Local Similarity 51.5%; Pred. No. 0.87;
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 238 GGGAGTGTGCGGTGCCCGGAGCGCGCGGCGGAGGCAAAAGGACGGAGCGGC 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17600 CCGCCGAGCCCCCGCGCCGCAACGCGGCGGAGAGAGAGCAAGCAACGCGTCG 17541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 GCGAGCGGGGCGGAGGTGCGTCGCCCTCCGCGCCCGCAGCGGTCGTCGCGTA 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17540 AGGTGCGGTCGCCGCGTAAGCGTCGACAGTTGCGGACACCGACCTTCGCCGTTG 17481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 358 GAGCGCAGCGCGCGCGATGAGCGGTGACCCGTCGCCCGCTCGGCGCGCAAGCGG 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17480 GCCCGTTGATGCGGTCAACCGGTGAGGTCGTGACGAGTCCGTCGTGACGCGGCGG 17421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 418 CCGTCGGGCTGCGCGCGCGGAG 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 17420 ACGTGGACACGCGCGACGCGG 17397

RESULT 14

ID AAI84136 standard; CDNA; 375 BP.

XX AAI84136;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 4196.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 KM nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

DR P-PSDB; AA004205.

PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

PS Claim 1; SEQ ID NO 4196; 1399pp + Sequence listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 375 BP; 54 A; 138 C; 113 G; 41 T; 29 other;

Query Match 9.1%; Score 45.4; DB 22; Length 375;
 Best Local Similarity 49.0%; Pred. No. 0.84;
 Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 292 CCGCGCGGACGCGGCGGAGCTTGCCTCCCTGCTGCGCCCGACGGGTTGCT 351
 DB 154 CCNCCNCCNNNNNNNCCGACNCCGCCCGGCTCCCGCCCGCGGCGGCGCGC 213
 QY 352 CGCGTAGAGCGGCGGCGCGGATGAAGCGGCTGAGCGCGGCGCGCTGCGGCGC 411
 DB 214 CGCGCGGCGCGGCGGCGCGGCGCGGCGCGGCGCGCGCGCGCGCGCGCG 273
 QY 412 AAGCGCGGCTGCGGCGGCGGCGGAGCTGCGCTGCGCTGCGCGGACGCGG 471

DB 274 CCCCNCNCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 333

QY 472 CACAGCCTGGGTGGCTCCGACGC 495

DB 334 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 357

RESULT 15

ID AAA31440 standard; DNA; 421 BP.

XX AAA31440;

DT 05-JUL-2000 (first entry)

DE Plant microsatellite marker #401.

XX Plant microsatellite sequence; core repeat sequence; detection; probe;
 KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;
 KM variety identification; genetic variability evaluation; primer; ss.

OS Eucalyptus grandis.

PN MO9967421-A1.

PD 29-DEC-1999.

PF 25-JUN-1999; 99MO-NZ00092.

PR 25-JUN-1998; 98US-0105307.

PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Havukkala IJ, Bloksberg LN, Glenn M;

XX WPI: 2000-116958/10.

PT New plant microsatellite markers and associated flanking species for
 PT the detection of polymorphic genetic markers -

PS Claim 1; Page 195; 392pp; English.

XX Sequences AAA31040-A32093 represent novel plant microsatellite sequences
 CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms, in genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for
 CC oligonucleotide fingerprinting and library screening and to design
 CC primers for microsatellite-primed PCR. Microsatellite markers are
 CC useful to locate specific economically useful genes in plant genomes.

XX Sequence 421 BP; 93 A; 142 C; 129 G; 57 T; 0 other;

Query Match 9.0%; Score 45.2; DB 21; Length 421;
 Best Local Similarity 52.1%; Pred. No. 0.92;
 Matches 101; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 296 CCGCGGACGCGGCGGAGCTTGCCTCCCTGCTGCGCCCGACGGGTTGCTGCG 355
 DB 228 CCGCGCGGCGGCTCTCTCTCTATCTCGGCGCTGCGCGGAGGCTTCTGCGC 287
 QY 356 TAGAGCGGAGGCGCGGCGGATGAAGCGGCTGAGCGCGGCGCGCTGCGGCGCAAG 415
 DB 288 TCGCGCTTCCCGCGGCGGAGTACGAGGACGCGCGCAAGCGCTGCGGCGCGCGG 347
 QY 416 CGCGCTGCGGCTGCGGCGGCGGAGCTGCGCTGCGCTGCGGCGGACGCGGCGACA 475
 DB 348 CGCGCGGAGTCCGCTGCGAGGCGGTGATAGCGGAGTCTGCGGCGGCGGCGAGC 407

OY 476 GCCTGGGTGGCTCC 489
| | | | | | |
Db 408 GCGAGGGGAGTCC 421

Search completed: October 10, 2002, 14:54:20
Job time : 287 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 18.7333 Seconds
(without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101A-5_COPY_1_500

Perfect score: 500
Sequence: 1 atgctcctgagcgcggtcc.....ccggcctgcagcagca 500

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCBUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.2	22.6	3138	US-09-234-332-5	Sequence 5, Appl1
2	56	11.2	4403765	US-09-103-840A-2	Sequence 2, Appl1
3	55.4	11.1	44377	US-08-804-227C-7	Sequence 7, Appl1
4	55.4	11.1	44377	US-08-804-198-1	Sequence 1, Appl1
5	51.4	10.3	1910	PCT-US92-05532-1	Sequence 1, Appl1
6	51.4	10.3	2261	US-08-272-882B-1	Sequence 1, Appl1
7	51	10.2	2370	US-08-104-072B-7	Sequence 7, Appl1
8	51	10.2	2370	US-08-351-413-8	Sequence 8, Appl1
9	51	10.2	2370	US-09-025-583-8	Sequence 8, Appl1
10	51	10.2	5173	US-08-242-677-1	Sequence 1, Appl1
11	50.6	10.1	23673	US-09-773-816-1	Sequence 1, Appl1
12	50.4	10.1	28804	US-08-592-874-1	Sequence 1, Appl1
13	50.4	10.1	28804	US-09-096-942-2	Sequence 2, Appl1
14	50.4	10.1	28804	US-09-096-867-2	Sequence 2, Appl1
15	49.8	10.0	423	US-09-144-085-5	Sequence 5, Appl1
16	49.8	10.0	8438	US-07-945-283-1	Sequence 1, Appl1
17	49.8	10.0	13842	US-09-105-537-30	Sequence 30, Appl1
18	49.8	10.0	36778	US-09-105-537-5	Sequence 5, Appl1
19	49.8	10.0	38506	US-09-320-878-19	Sequence 19, Appl1
20	49	9.8	4257	US-09-259-821A-1	Sequence 1, Appl1
21	49	9.8	4257	US-08-843-659-1	Sequence 1, Appl1
22	49	9.8	12001	US-08-458-568A-11	Sequence 11, Appl1
23	49	9.8	13987	US-08-804-227C-13	Sequence 13, Appl1
24	49	9.8	44377	US-08-804-227C-7	Sequence 7, Appl1
25	49	9.8	44377	US-08-804-198-1	Sequence 1, Appl1
26	49	9.8	44377	US-08-804-198-1	Sequence 1, Appl1
27	48.6	9.7	4848	US-08-955-957A-1	Sequence 1, Appl1

28	48.6	9.7	4848	US-08-955-957A-4	Sequence 4, Appl1
29	48.6	9.7	4848	US-08-955-957A-6	Sequence 6, Appl1
30	48.4	9.7	1578	US-08-681-129-1	Sequence 1, Appl1
31	48.2	9.6	2721	5215881-2	Patent No. 5215881
32	48	9.6	1203	US-09-086-010-1	Sequence 1, Appl1
33	48	9.6	3836	US-08-216-260-1	Sequence 1, Appl1
34	48	9.6	4257	US-08-690-473-1	Sequence 1, Appl1
35	48	9.6	4257	US-09-259-821A-1	Sequence 1, Appl1
36	48	9.6	4257	US-08-843-659-1	Sequence 1, Appl1
37	48	9.6	12001	US-08-458-568A-11	Sequence 11, Appl1
38	47.8	9.6	1656	US-09-385-028-14	Sequence 14, Appl1
39	47.8	9.6	15079	US-09-385-028-1	Sequence 1, Appl1
40	47.6	9.5	1153	US-09-372-448A-5	Sequence 5, Appl1
41	47.6	9.5	2639	US-07-952-817-8	Sequence 8, Appl1
42	47.6	9.5	2639	5210025-1	Patent No. 5210025
43	47.6	9.5	3181	US-09-135-021-1	Sequence 1, Appl1
44	47.6	9.5	3181	US-09-135-020-1	Sequence 1, Appl1
45	47.6	9.5	3181	US-09-135-010A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-234-332-5
; Sequence 5, Application US/09234332A
; Patent No. 6087168
; GENERAL INFORMATION:
; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michael F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3138
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: gene
; LOCATION: (0)...(0)
; OTHER INFORMATION: Z1c 1 Protein gene; Genbank Accession D76435
US-09-234-332-5

Query Match 22.6%; Score 113.2; DB 3; Length 3138;
Best Local Similarity 63.3%; Pred. No. 1.4e-13;
Matches 219; Conservative 0; Mismatches 103; Indels 24; Gaps 2;

QY	139	CGCCGCGCATGAGGAGCTTAACCTCAACCCGGGCGCGACGAGCTGTCCCGGCGCA	218
DB	900	CGCCGCGCATGAGGAGCTTAACCTCAACCCGGGCGCGACGAGCTGTCCCGGCGCG	959
QY	219	GAGCTCGGCGTTCAGCTGCGAGGCGCGCTTAACCCGGGCGCTGCGCTGCCGC	278
DB	960	CGAGACAGCTTCAGCTGCGAGGCGCGCTTAACCCGGGCGCTGCGCTGCCGC	1001
QY	279	TGCGGCGCGAGGCTTCGCGGCGCGCGCGCGAGCTGTGCTCTTGTGGCGCGCGG	338
DB	1002	TGCGGCGCGAGGCTTCGCGGCGCGCGCGCGAGCTGTGCTCTTGTGGCGCGCGG	1061
QY	339	CACCTTCACCGCGAGCTTCGCGGCGCGCGCGAGCTGTGCTCTTGTGGCGCGCGG	398
DB	1062	CACCTTCACCGCGAGCTTCGCGGCGCGCGCGAGCTGTGCTCTTGTGGCGCGCGG	1121
QY	399	CGCGGCGCGAGGCTTCGCGGCGCGCGCGAGCTGTGCTCTTGTGGCGCGCGG	452
DB	1122	CGCGGCGCGAGGCTTCGCGGCGCGCGCGAGCTGTGCTCTTGTGGCGCGCGG	1181
QY	453	CTCGGAGCGCGAGGCGCGAGCTTCCTTCCCGGCGCTGCGAGGAG	498


```
FEATURE:
NAME/KEY: promoter
LOCATION: 1..1808 /function= "another specific PT42
OTHER INFORMATION: promoter"
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1748..1755
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1780
OTHER INFORMATION: /product= "transcription
OTHER INFORMATION: initiation"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1809
OTHER INFORMATION: /product= "ATG start translation
OTHER INFORMATION: T42"
US-08-104-072B-7
```

```
Query Match 10.2% Score 51; DB 1; Length 2370;
Best Local Similarity 48.7%; Pred. No. 0.051;
Matches 167; Conservative 0; Mismatches 175; Indels 1; Gaps 1;
```

```
OY 7 CTGAGCGGGGTCGCGATTCCTCCGCGCATCGGGGTGGGCGACGTTCCGGCGCCGACCATCAC 66
DB 1799 CTAGACACCATGACACAGATCGATGCGCTCTCAAGGCGGTGGCGCCCTCTCTCTCATCT 1858
OY 67 CACTC-CGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 125
DB 1859 CATGCTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1918
OY 126 GCGGCGGGGGCGAGACGGCTTCGTTATTCGCGCGCGCGCGACATGGAGCCCTTCAAGCT 185
DB 1919 GCTGATGGGGCCCTGATGACCTGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 1978
OY 186 CAACCGGGGGGGCGACAGCTGTCCCGGGCGAGAGCTGGCGGCTTCAAGCTGCGAGGGGCC 245
DB 1979 CTGCTGCGAGAGCTGCGCGCTGCTCAACCGGATGGCTCCCGCGCGCGCGCGCGCGCG 2038
OY 246 GCGCGCTTACCCCGGGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305
DB 2039 CGTCTGCAATGGGTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2098
OY 306 CGCGCACGTTGGCTCTGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 348
DB 2099 CTCGCGCCCTCCCGCGCGCGCTGCGCGCGCTCTCATCAGCTTCAAC 2141
```

```
RESULT 8
US-08-351-413-8
Sequence 8, Application US/08351413
Patent No. 5750867
GENERAL INFORMATION:
APPLICANT: Williams, Mark
TITLE OF INVENTION: Maintenance of male-sterile plants
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 2046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/351.413
FILING DATE:
```

```
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/899,072
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,849
FILING DATE: 03-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-102PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2370 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Oryza sativa
STRAIN: Akihikari
FEATURE:
NAME/KEY: -
LOCATION: 1..1808 /label= PT42
OTHER INFORMATION: /note= "sequence comprising another specific
OTHER INFORMATION: promoter PT42"
FEATURE:
NAME/KEY: -
LOCATION: 1748..1755
OTHER INFORMATION: /label= TATA
OTHER INFORMATION: /note= "TATA Box"
FEATURE:
NAME/KEY: -
LOCATION: 1780
OTHER INFORMATION: /note= "transcription initiation
OTHER INFORMATION: site determined by primer extension"
FEATURE:
NAME/KEY: -
LOCATION: 1809
OTHER INFORMATION: /label= ATG
OTHER INFORMATION: /note= "ATG start of translation of rice T42 gene"
US-08-351-413-8
Query Match 10.2% Score 51; DB 1; Length 2370;
Best Local Similarity 48.7%; Pred. No. 0.051;
Matches 167; Conservative 0; Mismatches 175; Indels 1; Gaps 1;
```

```
OY 7 CTGAGCGGGGTCGCGATTCCTCCGCGCATCGGGGTGGGCGACGTTCCGGCGCCGACCATCAC 66
DB 1799 CTAGACACCATGACACAGATCGATGCGCTCTCAAGGCGGTGGCGCCCTCTCTCTCATCT 1858
OY 67 CACTC-CGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 125
DB 1859 CATGCTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1918
OY 126 GCGGCGGGGGCGAGACGGCTTCGTTATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
DB 1919 GCTGATGGGGCCCTGATGACCTGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 1978
OY 186 CAACCGGGGGGGCGACAGCTGTCCCGGGCGAGAGCTGGCGGCTTCAAGCTGCGAGGGGCC 245
DB 1979 CTGCTGCGAGAGCTGCGCGCTGCTCAACCGGATGGCTCCCGCGCGCGCGCGCGCGCG 2038
OY 246 CGCGCGCTTACCCCGGGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305
DB 2039 CGTCTGCAATGGGTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2098
```



```

SEQUENCE CHARACTERISTICS:
LENGTH: 5173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4863
US-08-242-677-1

Query Match      10.2%; Score 51; DB 1; Length 5173;
Best Local Similarity 48.0%; Pred. No. 0.05; Mismatches 220; Indels 6; Gaps 2.
Matches 209; Conservative 0;

QY   53 CGCGGCACCATTCACCACATCCGCGCGCGGGCGGGCGGAGCTCCGCCGAGATGACGAGACC 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   779 CGGGGTCTCGGGCCCCGCTTCCGCGCGCCGCGCGCGGCGGTGCGCCCGCGGACTTCGGGCAAC 720
QY   113 GTGAACTAGCGCTTGCGGGCGCGGCGCAGAAGCGCTTCGTTGATTCCGCCGCCGCGCACATGG 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   719 AGCTCTCGCGCAGGGCGCTCAGAGCACAGACCTTCTCTTACGCGCGCGGACCCACAG 660
QY   173 GAGCTTTAAGCTCAAACCCGGGCGGCGCAAGACTGTCCCAGGCGCAGACTTCGGGCGTTCA 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   659 GACGCCCCAGGCGCGCGCCAGCCGCCACACAGAGCCCGCAGCGCCGCCGCCACACTTGG 600
QY   233 GCTCCAGAGGCCCGCGCGCTCAACCCCGGCTCCGTGCGAGCTCCGCTGCGGCGCGACGCC 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   599 ACCACACACTG--GCAAGCAGTGCGCCCGGCGCACAGCGCCCGCGTCTGCGAGGCCCGC 542
QY   293 TCGGGCCCCAGCCCGCGACGTTGGCTCTACTGTGGCGGCGGCGCTTCAACTCCACCGGG 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   541 CCTCATCCCGCTCCCGCGCCCGCCCGCCAGCGCAGGCGAGAGGGGTCCCGCCACCCGC- 483
QY   353 ACTTCCTGTTCCGACAGCGCGCGGCTTCCGGGCACTTCGCGCGGCGGGCGGCGGACGACG 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   482 ---TCCAGTAGGGGCGCCGTCGTGCGGGGCGCGGAACATGGGCCGACGGCTGTACACT 426
QY   413 GCCTGTTCCGGCGCGGCGCGGCGGCGGCTTCCAGCAGCGCAGCTGGAGCGCCAGGGCCACC 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   425 TTCACGCGCAGCTTCGGCGCGCAGGCGGCGCACCCGCGCAGACATGCGCCAGCGCTCC 366
QY   473 TCCTCTTCCGGGCGC 487
    || | | | | | | | | |
DB   365 TCAGCCAGCGCGGCGC 351

RESULT 11
US-09-773-816-1
Sequence 1, Application US/09773816
Patent No. 6340774
GENERAL INFORMATION:
APPLICANT: Stanford University
APPLICANT: Khosla, Chaitan
TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
TITLE OF INVENTION: ANTAGONISTS
FILE REFERENCE: 28600-20210.00
CURRENT APPLICATION NUMBER: US/09/773, 816
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/243,458
PRIOR FILING DATE: 2000-10-25
PRIOR APPLICATION NUMBER: US 60/179,305
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 23673
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(23623)

```

[illegible]

RESULT 12
US-08-592-874-1
Sequence 1, Application US/08592874
Patent No. 5854034
GENERAL INFORMATION:
APPLICANT: POLLOCK, THOMAS J.
APPLICANT: YAMAZAKI, MOTOHIDE
APPLICANT: THORNE, LINDA
APPLICANT: MIKOLAJCZAK, MARCIA
APPLICANT: ARMENTROUT, RICHARD W.
TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
TITLE OF INVENTION: POLYSACCHARIDE PRODUCTION
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: JULES E. GOLDBERG
STREET: 261 MADISON AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,874
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/377,440
FILING DATE: 24-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: GOLDBERG, JULES E.
REGISTRATION NUMBER: 24,408
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-986-4090

TELEFAX: 212-818-9479
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28804 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 FRAGMENT TYPE: N-terminal
 US-08-592-874-1

Query Match 10.1%; Score 50.4; DB 2; Length 28804;
 Best Local Similarity 47.7%; Pred. No. 0.061;
 Matches 147; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 182 AGCTCAACCGGGGCGGCGACAGAGCTGTCGGGGCCAGAGCTGCGGCTTCAGCTGCAGG 241
 DB 14456 ATCTCGATACGAGAGAGCTGCGACAGCTCCAGGGCGGAGATACGCTCGGACCGG 14515
 QY 242 GCCCGGCGCTTACCCCGCTCGCTGCGGCTCGGCGCGACGCCCTCGGGCCCC 301
 DB 14516 TCGTCGCGCGCGCGCGCGCGATGGCGCGCGAGCTCCAGCTTCAGATGACGACG 14575
 QY 302 ACGCCGCGCGAGCTTGGCTTCTACTCTGGGCGCGCGCTTCACACTCCGCGAGCTTCTGT 361
 DB 14576 AGATCCAGAGCGGCGCGCGCGCAATCTGGCGCGCGCGCGCGCGCGCTTCAGGCGCTGACG 14635
 QY 362 TCCGCAAGCGCGCGCTTCCGGGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCTTTCG 421
 DB 14636 GCGAGCGCGGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 14695
 QY 422 GCG 481
 DB 14696 GCG 14755
 QY 482 CGGCGCTG 489
 DB 14756 ACCGCCAG 14763

RESULT 13

US-09-096-942-2
 Sequence 2, Application US/09096942
 Patent No. 6027925

GENERAL INFORMATION:
 APPLICANT: Pollock, Thomas J
 APPLICANT: Mikolajczak, Marcia
 APPLICANT: Yamazaki, Motohide
 APPLICANT: Thorne, Linda
 APPLICANT: Armentrout, Richard W
 TITLE OF INVENTION: Production of Xanthan Gum by *Sphingomonas* Bacteria
 FILE REFERENCE: seq list for appl filed from pro. appl
 CURRENT APPLICATION NUMBER: US/09/096,942
 EARLIER FILING DATE: 1998-06-12
 EARLIER APPLICATION NUMBER: 60/049,428
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 2
 LENGTH: 28804
 TYPE: DNA
 ORGANISM: *Sphingomonas* sp. S88
 US-09-096-942-2

Query Match 10.1%; Score 50.4; DB 3; Length 28804;
 Best Local Similarity 47.7%; Pred. No. 0.061;
 Matches 147; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 182 AGCTCAACCGGGGCGGCGACAGAGCTTCCCGGCGCGAGAGCTGCGGCTTCAGCTGCAGG 241
 DB 14456 ATCTCGATACGAGAGAGCTGCGACAGCTCCAGGGCGGAGATACGCTCGGACCGG 14515

QY 242 GCCCGGCGCTTACCCCGCTCGCTGCGGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
 DB 14516 TCGTCGCGCGCGCGCGCGCGATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 14575
 QY 302 ACGCCGCGCGAGCTTGGCTTCTACTCTGGGCGCGCGCTTCACACTCCAGCGGAGCTTCTGT 361
 DB 14576 AGATCCAGAGCGGCGCGCGCGCAATCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 14635
 QY 362 TCCGCAAGCGCGCGCTTCCGGGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 421
 DB 14636 GCGAGCGCGGAGCTTGGCTTCTACTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 14695
 QY 422 GCGCGGCGCGCGCGCGCGCTTCACAGCGCGAGCTGCGAGCGCGCGCGCGCGCGCGCG 481
 DB 14696 GCGCGGCGCGCGCGCGCGCTTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 14755
 QY 482 CGGCGCTG 489
 DB 14756 ACCGCCAG 14763

RESULT 14

US-09-096-867-2
 Sequence 2, Application US/09096867
 Patent No. 6030817

GENERAL INFORMATION:
 APPLICANT: Pollock, Thomas J
 APPLICANT: Mikolajczak, Marcia
 APPLICANT: Yamazaki, Motohide
 APPLICANT: Thorne, Linda
 APPLICANT: Armentrout, Richard W
 TITLE OF INVENTION: Production of Xanthan Gum by *Sphingomonas* Bacteria
 FILE REFERENCE: seq list for appl filed from pro. appl
 CURRENT APPLICATION NUMBER: US/09/096,867
 EARLIER FILING DATE: 1998-06-11
 EARLIER APPLICATION NUMBER: 60/049,428
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 2
 LENGTH: 28804
 TYPE: DNA
 ORGANISM: *Sphingomonas* sp. S88
 US-09-096-867-2

Query Match 10.1%; Score 50.4; DB 3; Length 28804;
 Best Local Similarity 47.7%; Pred. No. 0.061;
 Matches 147; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 182 AGCTCAACCGGGGCGGCGACAGAGCTTCCCGGCGCGAGAGCTGCGGCTTCAGCTGCAGG 241
 DB 14456 ATCTCGATACGAGAGAGCTGCGACAGCTCCAGGGCGGAGATACGCTCGGACCGG 14515
 QY 242 GCCCGGCGCTTACCCCGCTCGCTGCGGCTCGGCGCGACGCCCTCGGGCCCC 301
 DB 14516 TCGTCGCGCGCGCGCGCGCGATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 14575
 QY 302 ACGCCGCGCGAGCTTGGCTTCTACTCTGGGCGCGCGCTTCACACTCCAGCGGAGCTTCTGT 361
 DB 14576 AGATCCAGAGCGGCGCGCGCGCAATCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 14635
 QY 362 TCCGCAAGCGCGCGCTTCCGGGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 421
 DB 14636 GCGAGCGCGGAGCTTGGCTTCTACTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 14695
 QY 422 GCGCGGCGCGCGCGCGCGCTTCACAGCGCGAGCTGCGAGCGCGCGCGCGCGCGCGCG 481
 DB 14696 GCGCGGCGCGCGCGCGCGCTTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 14755
 QY 482 CGGCGCTG 489
 DB 14756 ACCGCCAG 14763

RESULT 15

US-09-144-085-5

; Sequence 5, Application US/09144085

; Patent No. 6280999

; GENERAL INFORMATION:

; APPLICANT: Gustafsson, Claes

; APPLICANT: Beliaich, Mary C.

; APPLICANT: Ashley, Gary

; APPLICANT: Julien, Bryan

; APPLICANT: Ziermann, Rainer

; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA

; TITLE OF INVENTION: THERFOR

; FILE REFERENCE: 30062-20020.20

; CURRENT APPLICATION NUMBER: US/09/144,085

; CURRENT FILING DATE: 1998-08-31

; EARLIER APPLICATION NUMBER: 09/010,809

; EARLIER FILING DATE: 1998-01-22

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 423

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: DNA fragment

; OTHER INFORMATION: corresponding to a KS domain of Sorangium

; OTHER INFORMATION: cellosporium SMP44 gene

US-09-144-085-5

Query Match 10.0%; Score 49.8; DB 4; Length 423;

Best Local Similarity 47.6%; Pred. No. 0.09;

Matches 147; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

```
QY 1 ATGCTCTGGACGCGGCTCCGACGATTCGCGCCATGCGGCGGAGAGCTTCGCGCGCCAC 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 ATGTTGTGTGAGAGTTCAAGCCGTACGCGGCGCTGCGCCGCGACGCGCAGCAAGGCTTC 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CATCACCACTCCGCGCGCGCGCGCGCTGCGCGCGAGATGACAGACCGTGAACGTG 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 TCGGCGGACGCGCGAGGTTCTGGCGCGGAGGCGGCTGCTGTCTGTGGAGCGG 234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 AGCTTGCGGCGCGCGAGACGCGCTTCGTGATTCGCGCCCGCGGACATGGAGCCTTC 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 CTCTCGGACGCGCGGCGCAACGGTCAACCGGTCTCGCTGCTGCGGTACCGCCGTC 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 AAGCTCAACCGGCGCGCGAGAGCTGCCGCGGCGAGACTGCGGCTTCACGTCCGAG 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 AACCGAGACGCGCGCAGCAACGGGCTTACCGCGCCCAACGAGACCTTCGCAAGCGGGTG 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 GGGCCCGCGCGCTACCGCGCTCGGTGCGGCTGCGCGCTCGCGCGCGAGCGCGCTCGGCCC 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 ATCCGGAGCGGCTGCGCGAGCGCGCGGTGCGACCGCGGACATCGACGCGCTCGAGACG 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 CACGCGCGG 309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 CACGCGACG 423
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 10, 2002, 17:52:13
Job time : 1157.73 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 Search time 695.733 Seconds
(without alignments)
9693.805 Million cell updates/sec

Title: US-09-489-101a-5_COPY_1_500

Perfect score: 500

Sequence: 1 atgctcctgagcagcggtcc.....ccggcctgcagagcagca 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	269	53.8	642	9	BB655427
2	224.2	44.8	278	9	BB655427
3	190.2	38.0	503	9	AT511335
4	115.7	31.4	496	9	AT511335
5	113.2	22.6	487	10	BE097311
6	108.4	21.7	441	9	BE097311
7	90	18.0	455	10	BE097311
8	86.8	17.4	935	12	BE097311
9	83.6	16.7	935	12	BE097311
10	80.2	16.4	925	12	BE097311
11	80.4	15.7	925	12	BE097311
12	78.4	15.4	821	12	BE097311
13	77.2	15.4	1369	12	BE097311
14	74.2	14.8	932	12	BE097311
15	73.6	14.7	976	12	BE097311
16	72.8	14.6	1008	12	BE097311
17	71.6	14.3	935	12	BE097311

C 18	71.6	14.3	1203	12	CNS01514	AL106054 Drosophila
C 19	71.2	14.2	957	12	AG077010	AG077010 Pan trogl
C 20	71	14.2	885	12	AG159162	AG159162 Pan trogl
C 21	70.4	14.1	693	12	AG166257	AG166257 Pan trogl
C 22	70.2	14.0	652	10	BJ057546	BJ057546 BJO57546
C 23	70.2	14.0	872	12	CNS0241R	AL188712 Tetradon
C 24	70	14.0	810	12	AG060267	AG060267 Pan trogl
C 25	70	14.0	841	12	A2196416	A2196416 SP.1031_B
C 26	70	14.0	992	12	AG057280	AG057280 Pan trogl
C 27	70	14.0	1021	12	AG061741	AG061741 Pan trogl
C 28	69.8	14.0	894	12	AG075627	AG075627 Pan trogl
C 29	69.6	13.9	813	12	A2193936	A2193936 SP.1025_A
C 30	69.4	13.9	1049	12	AG185368	AG185368 Pan trogl
C 31	69.2	13.8	894	12	AG075627	AG075627 Pan trogl
C 32	69.2	13.8	1452	12	AG032979	AG032979 Pan trogl
C 33	69	13.8	1094	12	AG036373	AG036373 Pan trogl
C 34	69	13.8	1160	12	AG043473	AG043473 Pan trogl
C 35	68.6	13.7	921	12	A2211117	A2211117 SP.0159_B
C 36	68.6	13.7	957	12	AG077010	AG077010 Pan trogl
C 37	68.6	13.7	1023	12	AG060164	AG060164 Pan trogl
C 38	68.6	13.7	1375	9	AW727483	AW727483 GA_Fa001
C 39	68.4	13.7	783	12	AG060185	AG060185 Pan trogl
C 40	68.4	13.7	1084	12	AG058435	AG058435 Pan trogl
C 41	68	13.6	663	12	AG080425	AG080425 Pan trogl
C 42	67.8	13.6	796	9	AU172073	AU172073 AU172073
C 43	67.8	13.6	685	10	BG786336	BG786336 SEAMC006
C 44	67.8	13.6	1131	12	AG042920	AG042920 Pan trogl
C 45	67.6	13.5	1165	12	AG030649	AG030649 Pan trogl

ALIGNMENTS

RESULT 1
BB655427
LOCUS
DEFINITION
CDNA clone D030057J08 5', mRNA sequence.
ACCESSION
BB655427
KEYWORDS
VERBOS
SOURCE
ORGANISM
Mus musculus
house mouse
EST.

REFERENCE
AUTHORS
Atakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Haru, A., Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Koda, M., Koyama, S., Matsuyama, T., Miyazaki, A., Momura, K., Ohno, H., Sasaki, O., Shibata, K., Shindagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyka, Y., Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Atakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
url:http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wang, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabe, K., Yoneda, Y., Ishikawa, T., Oawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kito, A., and Hayashizaki, Y.

RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, T., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*, 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES	Location/Qualifiers
source	1, .642

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="D030057J08"
 /clone_1ib="RIKEN full-length enriched, 9 days embryo"
 /dev_stage="9 days embryo"
 /lab_host="DH10B"
 /note="Site.1: SalI; site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGACAGACAGCGCCGCCACACAGATTTTCTTTTCTTTTNN 3'], cDNA was prepared by using tetraolase thermo activated reverse transcriptase and subsequently enriched for full-length by cap trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
 GAGACAGAGATTTTCGATTTATTTATATTCACCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluscript KSI(+) after bulk excision from lambda FLIC I."

BASE COUNT
ORIGIN

Query Match	53.8%	Score 269	DB 9	Length 642
Best Local Similarity	85.4%	Pred. No. 6,3e-30		
Matches 299	Conservative	0	Mismatches 51	Indels 0
			Gaps	0
QY	1	ATGCTCTGAGACGGGGGTCCGACAGTTCCCGGACATCGGGGTGGGCACTTCGGCGCCAC	60	
DB	293	ATGCTTTTGGACCGGGGCGCGAGTTCCCGGCAATCGGGGTGGGCACTTCGGCGCCAC	352	
QY	61	CATCACCACTCCCGCCGCGGCGGCGGCGGCGTGCACCGCGAATGCAGGACCTGTGACATG	120	
DB	353	CACCACTCACTCGCGCCCGGCGACGCGCGCGCTGGCGGCGCCGAGATGCAGGACCCGAGATG	412	
QY	121	AGCTGGCGGCGCGCGGAGAAAGGCTTCGTTGATTCCCGCGCGCGGCGCATGGAGGCTTC	180	
DB	413	AGCTGGCGGCGCACTAAACGCGCTTCGATGCACTGGCGCGCGGCGCATGGGCGCTTC	472	
QY	181	AACTCAACCCGGGCGCGCAAGCTGTCCCGGGCCAGAGCTTGCGCTTACGTGCAG	240	
DB	473	AAAGTCAAACCCCGGGCGACAGAACTGTCTCTGTGTAGAGTTGGGGTTCACTGCGCAA	532	
QY	241	GGCCCGGCGCTACACCCCGGCTCCCTCTCGGCTGGCGTGGCGCGGCAAGGCTCGGAGCC	300	
DB	533	GGTCCGGGTCTTACCCCGGGTTTCGGCTGCTGTGAGCGCTGGCGCGCGGTCTTAAAGGCC	592	
QY	301	CAGCGCGGCACTTGGCTCTACTCTGTGGCGGCCCTTCAACTCCACCG	350	
DB	593	CAGCGCGCACTGGCTCTATTCTGTGGCGCTTCACTTATTCACCGG	642	

RESULT 2
AI511335/C

LOCUS	A1511335	278 bp	mRNA	linear	EST 03-JUL-1999
DEFINITION	UI-R-C3-sf-a-03-0-UI.s1	UI-R-C3	Rattus norvegicus	cdna	clone
ACCESSION	UI-R-C3-sf-a-03-0-UI 3'				
VERSION	A1511335				
KEYWORDS	A1511335.1	GI:4417034			
SOURCE	EST.				
ORGANISM	NOTIFY RAT				
	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

1 (bases 1 to 278)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Address: Department of Molecular Biology and Genetics, University of

MEDLINE
COMMENT

Email: mscores@blue.weeg.illowa.edu
oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message. cdna library Preparation: M.B.
Scars lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) This clone is also available
through the I.M.A.G.E. Consortium at LNL (info@image.lnl.gov).
IMAGE ID=1768598
Seq primer: M13 Forward
POLYA=NO.

FEATURES
Source

38 a 102 c 98 g 40 t

BASE COUNT
ORIGIN

38 a 102 c 98 g 40 t

Research 6:791-806, 1996)

genome

and Soares,

Life technologies) and

to generate the ui-R-C3 library. This procedure has been

previously described (Bonaldo,

chromatography, converted to double-stranded circles and

electroporated into DH10B bacteria (Life technologies) to

library) was purified by hydroxyapatite column

circles. The remaining single-stranded circles (subtracted

the ui-R-C2p library in the form of single-stranded

been derived was used as a driver in a hybridization with

CNA inserts from ui-R-C2p clones from which 3' ESTs had

library (ui-R-C3) was constructed as follows: PCR amplified

of origin of a clone within the mixture. The subtracted

oligo-dt track which allows identification of the library

nucleotides present between the Not I site and the

ui-R-C0, ui-R-A1, ui-R-E1. The tag is a string of 3-5

generated. The following serially subtracted libraries

embryos, after a series of subtractions to reduce the

kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day

libraries from rat placenta, adult lung, brain, liver,

derived from a mixture of individually tagged normalized

library is a subtracted library of a series, ultimately

polylinker. Site 1: Not I; Site 2: Eco RI. The ui-R-C3

/note/vector: p733p-Pac (Pharmacia) with a modified

/lab.host="DH10B (Life technologies)"

/dev.stage="adult"

/clone.lib="ui-R-C3"

/clone="ui-R-C3-sf-a-03-0-UI"

/db.xref="Taxon:10116"

/db.xref="Sprague-Dawley"

/organism="Rattus norvegicus"

BASE COUNT
ORIGIN

Query Match	44.8%;	Score 224.2;	DB 9;	Length 278;
Best Local Similarity	88.1%;	Pred. No. 1.7e-23;		
Matches 24;	Conservative 0;	Mismatches 33;	Indels 0;	Gaps 0;

```

QY 1 ATGCTCTGAGCGGGGTTCGCGAGTTCCTCCGCGCATCGGGGTGGGCAAGCTTCGCGCGCCAC 60
Db 277 ATGCTCTGAGCGGGGTTCGCGAGTTCCTCCGCGCATCGGGGTGGGCAAGCTTCGCGCGCCAC 218
QY 61 CATACCACTCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120
Db 217 CACCACTCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 158
QY 121 AGCCTGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 180
Db 157 AGCCTGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 98
QY 181 AAGCTCAACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
Db 97 AAGCTCAACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 38
QY 241 GGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 277
Db 37 GGTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1

```

RESULT 3

```

AI359298 503 bp mRNA linear EST 15-FEB-1999
LOCUS qy29c02.x1 NCI-CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2013410 3'
DEFINITION similar to TR:062520 Q62520 ZIC2 PROBEIN.; contains element MSRL
repetitive element; , mRNA sequence.
ACCESSION AI359298
VERSION AI359298.1 GI:4110919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 503)
AUTHORS NCI/NIHDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index
Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bhrp/image/image.html
Insert Length: 1692 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 488.
Location/Qualifiers
1. 503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2013410"
/clone_lib="NCI-CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAAGTGGAGGCGCGGCATCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

```

BASE COUNT
ORIGIN

83 a 193 c 159 g 68 t

```

Query Match 38.0%; Score 190.2; DB 9; Length 503;
Best Local Similarity 97.3%; Pred No. 1.3e-18;
Matches 215; Conservative 0; Mismatches 3; Indels 3; Gaps 2;
QY 280 GCGGCGCAGCGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 339
Db 1 GCGGCGCAGCGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 60
QY 340 AACTCCACCGGGGACTTCTCTTCCGCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 399
Db 61 AACTCCACCGGGGACTTCTCTTCCGCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 117
QY 400 GGGGGGCGACGAGGGGCTGTTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 459
Db 118 GGGGGGCGACGAGGGGCTGTTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 177
QY 460 GCGCAGGGGCGACCTCTCTTCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 500
Db 178 GCGCAGGGGCGACCTCTCTTCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 218

```

RESULT 4

```

BE097311 496 bp mRNA linear EST 12-JUN-2000
LOCUS UI-R-B01-apv-h-02-0-UI.s1 UI-R-B01 Rattus norvegicus cDNA clone
DEFINITION UI-R-B01-apv-h-02-0-UI 3', mRNA sequence.
ACCESSION BE097311
VERSION BE097311.1 GI:8488242
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 496)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mscares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-No.
Location/Qualifiers
1. 496
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-B01-apv-h-02-0-UI"
/clone_lib="UI-R-B01"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The library
(UI-R-B01) is a subtracted library derived from a mixture
of the following tissues: thalamus, cerebellum,
hypothalamus, medulla, pons, midbrain, cerebral cortex,
corpus striatum and hippocampus. For a detailed
description of the library from which this clone was
derived, please visit our web site at
rategen.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research

```

FEATURES

source

FEATURES

source

6:791-806, 1996)
TAG_SEQ=None found"
BASE COUNT 82 a 179 c 140 g 95 t
ORIGIN

Query Match 31.4%; Score 157; DB 9; Length 496;
Best Local Similarity 83.4%; Pred. No. 7.3e-14;
Matches 191; Conservative 0; Mismatches 35; Indels 3; Gaps 1;

QY 272 CTGCGCGCTGGCGCGCGCTGCGCGCCACGCGCGGACGTTGGCTCCACTCTGGG 331
D 1 CGCGCGCTGAGCGCGCGCGCTCTAGGCGCCACGCGCTCAGCTTGCGCTTATTCGGGCG 60
QY 332 CGCGCTTCAACTCCACCGCGGACTTCTGTTCGCGACGCGCGCTTCCGGGACTTGG 391
D 61 CTCCTTTATTCACCGCGGACTTCTGTTCGCTAGC---CGGGGCTTGGGGACTCGG 117
QY 392 CGCGCGCGCGCGCGCGCGCTGTTCGGGCGCGCGCGCGCGCTTCCGACGACGCGCG 451
D 118 CGCGCGGAGCGCGCGCGCGCTTTCGACCGCGCGCGCGCGCTTCCACCATGCGC 177
QY 452 ACTCGAGCGCGACGCGCGCTCTCTTCCGGGCGCTGCGACGACGCA 500
D 178 ACTCGAGCGCGCGCGCGCTCTCTTCCGCGCTTCCGGAACGCA 226

RESULT 5
LOCUS BF441548 487 bp mRNA linear EST 01-DEC-2000
DEFINITION 257879 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF441548
VERSION BF441548.1 GI:11501640
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grose, W.M., Bennett, G.A., Laegreid, W.W., and Keeler, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smithemall.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross-match with the -minscore 18 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCACGAGC
Plate: 86 row: G column: 7
Seq primer: ATTAGGTGACATATAG.
FEATURES
Location/Qualifiers
1..487

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI; library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 69 a 192 c 152 g 74 t
ORIGIN

Query Match 22.6%; Score 113.2; DB 10; Length 487;
Best Local Similarity 63.3%; Pred. No. 1.4e-07;

Matches 219; Conservative 0; Mismatches 103; Indels 24; Gaps 2;

QY 159 CGCGCGCACATGGGAGCTTCAAGCTCAACCCGCGCGCGCGAGCTGTCCCGGGCCA 218
D 142 CGCGGACGATGCGCGCTTCAAGCTCAACCCGCGCGAGCTCTCAGAGCTGCTCCGGCG 201
QY 219 GAGCTGGGCTTCACTCCGAGGCGCGCGCGCTTACCGCGCTCCGCTGCGCGCG 278
D 202 CCAGACGCTTCACTCCGAGGCTCTTGGCTAC-----GCGCTGC 243
QY 279 TCGCGCGCGAGCGCTTGGGCGCGCGCGCGCGAGCTTGGCTCTACTGTGGCGCGCTT 338
D 244 GCGCGCGCTGGGACATCAGCATCAGCGCGCGCGCGCGCTTCTATTCAGCGCGAGCTT 303
QY 339 CAATCTCACCGGAGCTTCTGTTCGCGACGCGCGCGCTTCCGGGACTTCCGCGCGCG 398
D 304 CAATCTCACCGGAGCTTCTGTTCGCGAACGCGGCGCTTGGGAGCGCGCGCGAGC 363
QY 399 CGCGCGCGCGACAGCGCTTTCGGGCGCGCGCGCGCGCTT-----GCACACGCGCA 452
D 364 GCGCGCGCGACAGCTTCTGTTCGCTTTCGGCTTTCGGCTTTCGGGCGCGCGCGCGCA 423
QY 453 CTCGAGCGCGAGCGCGCTCTCTTCCGGGCGCTGCGACGACG 498
D 424 CACGAGCGCGCGCGCGCTCTTTCGGGCGCTTTCGCGACGACG 469

RESULT 6
LOCUS A1894008 441 bp mRNA linear EST 15-MAR-2000
DEFINITION mg08e03.y1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA clone IMAGE:423196 5', similar to gb:U32167 mouse mRNA for zic protein (MOUSE);, mRNA sequence.
ACCESSION A1894008
VERSION A1894008.1 GI:5599910
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 441)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Waller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: mg08e03.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the correct orientation)

Seq primer: -40RP from Glibco
High quality sequence stop: 413.
Location/Qualifiers
1..441

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="IMAGE:423196"
/clone_lib="Soares mouse embryo NBME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"

[illegible]

BASE COUN
ORIGIN

BASE COUNT	45 a	382 c	553 g	32 t	357 others
ORIGIN	/cell_type="lymphoblastoid",clone_id="PTB Chimpanzee Male BAC Library"				
Query Match	15.4%	Score 77.2;	DB 12;	Length 1369	
Best local Similarity	47.2%;	Pred. No. 0.019;			

[illegible][illegible]

RESULT 15
AG041117/c
LOCUS
AG041117
ACCESSION
AG041117
VERSION
AG041117.1 GI:16569842
KEYWORDS
GSS; GSS (genome survey sequence).
Pan troglodytes male lymphoblast DNA, clone_11b:PTB Chimpanzee Male
SOURCE
BAC Library clone:PTB-018L17.F.
Pan troglodytes
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 (sites)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of library PTB
Unpublished
2 (bases 1 to 976)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suenho-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the Rad process and may have higher chance of
clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : PKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .976

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/clone="PTB-018L17.F"

/sex="male"

/cell_type="Lymphoblast"

/clone_11b="PTB Chimpanzee Male BAC Library"

BASE COUNT

59 a 287 c 495 g 51 t 84 others

ORIGIN

Query Match 14.7%; Score 73.6; DB 12; Length 976;

Best Local Similarity 48.2%; Pred. No. 0.064;

Matches 238; Conservative 0; Mismatches 245; Indels 11; Gaps 1;

```
QY 6 CCTGGAGCCGGGTCGCGAGTTCCGGCCATCGGGGTGGGCAAGCTTCGGCCGCCACATCA 65
    || || || || || || || || || || || || || || || || || || || || ||
Db 835 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 776
    || || || || || || || || || || || || || || || || || || || || ||
QY 66 CCACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
    || || || || || || || || || || || || || || || || || || || || ||
Db 775 NCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 716
    || || || || || || || || || || || || || || || || || || || || ||
QY 126 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
    || || || || || || || || || || || || || || || || || || || || ||
Db 715 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 656
    || || || || || || || || || || || || || || || || || || || || ||
QY 186 CAACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 245
    || || || || || || || || || || || || || || || || || || || || ||
Db 655 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 596
    || || || || || || || || || || || || || || || || || || || || ||
QY 246 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 305
    || || || || || || || || || || || || || || || || || || || || ||
Db 595 G-----CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 547
    || || || || || || || || || || || || || || || || || || || || ||
QY 306 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 365
    || || || || || || || || || || || || || || || || || || || || ||
Db 546 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487
    || || || || || || || || || || || || || || || || || || || || ||
QY 366 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 425
    || || || || || || || || || || || || || || || || || || || || ||
Db 486 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 427
    || || || || || || || || || || || || || || || || || || || || ||
QY 426 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 485
    || || || || || || || || || || || || || || || || || || || || ||
Db 426 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
    || || || || || || || || || || || || || || || || || || || || ||
QY 486 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 499
    || || || || || || || || || || || || || || || || || || || || ||
Db 366 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 353
```

Search completed: October 10, 2002, 20:27:22
Job time : 701.733 secs

XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
 XX WPI: 2001-457597/49.
 DR P-PSDB: AAE05812.
 XX Isolated polypeptide, used to treat or prognose a disorder
 PT characterized by expression of a hCAP e.g. cancer, is encoded by an
 PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
 XX
 PS Claim 57; Page 90-91; 152pp; English.
 CC The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterised by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterised by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterised by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are small and non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
 CC The present sequence is a small cell lung cancer associated gene
 CC designated as NY-SCC-3 encoding human ZIC2 protein.
 XX
 SQ Sequence 1602 BP; 275 A; 578 C; 536 G; 213 T; 0 other;
 Query Match 100.0%; Score 500; DB 22; Length 1602;
 Best Local Similarity 100.0%; Pred. No. 3e-75;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ATGCTCCGAGAGGGGGGCGAGTCCGGCCATCCGGGGGTGGGAGCTTCCGGCCAC 60
 DB 1 ATGCTCCGAGAGGGGGGCGAGTCCGGCCATCCGGGGGTGGGAGCTTCCGGCCAC 60
 OY 61 CATCACACTCCG 120
 DB 61 CATCACACTCCG 120
 OY 121 AGCTGGGGGGGGGCGAGAGGGTGTGATTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 DB 121 AGCTGGGGGGGGGCGAGAGGGTGTGATTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 OY 181 AAGCTCAACCGGGGGCGCGAGAGCTGTCCCGGGCGCGAGCTCGCGTCACTCGCAG 240
 DB 181 AAGCTCAACCGGGGGCGCGAGAGCTGTCCCGGGCGCGAGCTCGCGTCACTCGCAG 240
 OY 241 GGGCGGGGCGCTACCG 300
 DB 241 GGGCGGGGCGCTACCG 300
 OY 301 CAGCGCGCGAGTGTGCTACTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 DB 301 CAGCGCGCGAGTGTGCTACTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 OY 361 TTCCGAGAGGGGGCGCTCCGGGAGCTTCCGGGAGCTTCCGGGAGCTTCCGGGAGCTT 420
 DB 361 TTCCGAGAGGGGGCGCTCCGGGAGCTTCCGGGAGCTTCCGGGAGCTTCCGGGAGCTT 420
 OY 421 GGGCGGGGGCG 480
 DB 421 GGGCGGGGGCG 480
 OY 481 CCGGGCGCGCGAGAGCA 500
 DB 481 CCGGGCGCGCGAGAGCA 500
 RESULT 2
 AAS61863
 ID AAS61863 standard; cDNA; 1602 BP.
 XX

AC AAS61863;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Lung small cell carcinoma antigen, cDNA #404.
 XX
 KW Human; cytosolic; antitumour; lung small cell cancer antigen;
 KW tumour; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200177168-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-US11859.
 XX
 PR 11-APR-2000; 2000US-196780P.
 PR 21-JUN-2000; 2000US-213361P.
 PR 01-SEP-2000; 2000US-229763P.
 PR 05-SEP-2000; 2000US-230629P.
 PR 14-SEP-2000; 2000US-232565P.
 PR 19-DEC-2000; 2000US-257037P.
 PR 08-JAN-2001; 2001US-260796P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Lodes MJ, Wang T, Mohamath R, Indirias CY;
 XX
 DR WPI: 2002-010896/01.
 DR P-PSDB: AAU69423.
 XX
 PT Lung tumour polynucleotide and polypeptides useful in therapy and
 PT diagnosis of cancer especially lung cancer -
 XX
 Claim 1; Page 274; 295pp; English.
 The invention relates to novel isolated lung small cell cancer antigen
 polynucleotides (I) and polypeptides (II) used in a method of detecting
 cancer in a patient. The method is optionally performed by
 utilizing oligonucleotides (III), where the biological sample
 from the patient is contacted with (III), detecting the amount of
 polynucleotide hybridised to (III) in the sample and comparing the
 amount of polynucleotide to a predetermined cut-off value and thereby
 determining cancer in a patient. (I), (II) or antigen-presenting cells
 expressing (II) is useful for stimulating and/or expanding T cells
 specific for a tumour protein. The method comprises contacting T cells
 with one of the components under conditions to permit the stimulation
 and/or expansion of the cells. A composition comprising (I) is useful for
 stimulating an immune response in a patient and for inhibiting the
 development of a cancer especially lung cancer in a patient. An
 isolated T cell population is useful for removing tumour cells from the
 biological sample and for inhibiting the development of cancer in a
 patient. AAS61460-AAS61874 represent novel human lung small cell
 cancer antigen coding sequences of the invention.
 XX
 SQ Sequence 1602 BP; 275 A; 578 C; 536 G; 213 T; 0 other;
 Query Match 100.0%; Score 500; DB 24; Length 1602;
 Best Local Similarity 100.0%; Pred. No. 3e-75;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ATGCTCCGAGAGGGGGGCGAGTCCGGCCATCCGGGGGTGGGAGCTTCCGGCCAC 60
 DB 1 ATGCTCCGAGAGGGGGGCGAGTCCGGCCATCCGGGGGTGGGAGCTTCCGGCCAC 60
 OY 61 CATCACACTCCG 120
 DB 61 CATCACACTCCG 120
 OY 121 AGCTGGGGGGGGGCGAGAGGGTGTGATTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 DB 121 AGCTGGGGGGGGGCGAGAGGGTGTGATTCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

```

QY 181 AAGCTCAACCCGGGCGGCGACAGCTCTCCCGGGCAGAGCTCGGCTTCACGTCGACG 240
DB 181 AAGCTCAACCCGGGCGGCGACAGCTCTCCCGGGCAGAGCTCGGCTTCACGTCGACG 240
QY 241 GGGCGGGGCGGCTACCCGGGCTCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300
DB 241 GGGCGGGGCGGCTACCCGGGCTCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 300
QY 301 CAGGCGGCGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
DB 301 CAGGCGGCGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 361 TTCCGCGACGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
DB 361 TTCCGCGACGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 421 GGGCGGGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
DB 421 GGGCGGGGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 481 CCGGCGCTGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 500
DB 481 CCGGCGCTGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 500

RESULT 3
AAA62683
ID AAA62683 standard; DNA: 3138 BP.
XX
AC AAA62683;
XX
DT 29-NOV-2000 (first entry)
XX
DE Human Zic1 gene.
XX
KW Human; Zic1; epidermal cell transdifferentiation; gene therapy;
KW cerebroprotective; neuroprotective; brain injury; spinal cord injury;
KW stroke; neurodegenerative disease; Parkinson's disease;
KW Huntington's disease; Alzheimer's disease; neuronal cell generation; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 781..2124
FT /tag= a
FT /product= "Zic1"
XX
PN US6087168-A.
XX
PD 11-JUL-2000.
XX
PE 20-JAN-1999; 99US-0234332.
XX
PR 20-JAN-1999; 99US-0234332.
XX
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI Levesque MF, Neuman T;
XX
DR WPI: 2000-498200/44.
DR P-PSDB: AAB14349.
XX
PT Converting epidermal cells into neurons, useful for isolating nerve
PT growth factors or for gene therapy, comprises dedifferentiating cells
PT and transfecting with vectors with a cDNA coding neurogenic
PT transcription factors
XX
PS Example 2; Column 23-26; 27pp; English.
XX
CC The present sequence is the human Zic1 gene from Genbank. It was
CC used to provide sequence information for the cloning of Zic1 cDNA,
CC which was used to transfect cultured epidermal cells. This was

```

```

CC part of a novel method for transdifferentiating an epidermal basal
CC cell into a cell having the morphological, physiological and/or
CC immunological features of a viable neuronal cell. The method is useful
CC for screening new drugs for treating a nervous system disorder, or for
CC isolating a novel nerve growth factor. The transdifferentiated cell is
CC useful in both cell and gene therapies aimed at alleviating various
CC neurological disorders. The cell or gene therapy approach involves the
CC use of autologous transplantation or grafting of the newly created
CC neuronal cells as treatment for brain or spinal cord injury, stroke and
CC neurodegenerative diseases (e.g. Parkinson's disease, Huntington's
CC disease or Alzheimer's disease).
XX
SQ Sequence 3138 BP; 701 A; 898 C; 840 G; 699 T; 0 other;
Query Match 22.6%; Score 113.2; DB 21; Length 3138;
Best Local Similarity 63.3%; Pred. No. 6,4e-11;
Matches 219; Conservative 0; Mismatches 103; Indels 24; Gaps 2;
QY 159 CGCGCGGCAATGGAGCCTTCAGCTCAACCGGCGCGGCGGCGGCGGCGGCGGCA 218
DB 900 CGCGGCGGCAATGGAGCCTTCAGCTCAACCGGCGGCGGCGGCGGCGGCGGCGG 959
QY 219 GAGCTCGGCTTCAGCTTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 278
DB 960 CAGACAGGCTTCAGCTTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1001
QY 279 TGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 338
DB 1002 TGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1061
QY 339 CAACCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 398
DB 1062 CAACCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1121
QY 399 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 452
DB 1122 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1181
QY 453 CTGGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 498
DB 1182 CACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1227

RESULT 4
AB199239
ID AB199239 standard; cDNA: 2947 BP.
XX
AC AB199239;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse
XX
PE Mouse ischaemic condition related cDNA sequence SEQ ID NO: 60.
XX
PR Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PE 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI: 2002-034733/04.
DR P-PSDB: ABB57042.
XX

```

Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes.

Claim 2: Page 191-194; 260pp; English.

The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI9202 to ABI9912, encoding the protein sequences in AB57020 to AB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI9913 and ABI9914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

Sequence 2947 BP; 659 A; 818 C; 787 G; 683 T; 0 other:

Query Match 21.0%; Score 105.2; DB 24; Length 2947;
Best Local Similarity 61.8%; Pred. No. 1.4e-09;
Matches 214; Conservative 0; Mismatches 108; Indels 24; Gaps 2;

```

159 CGCGGCGCAGTGGAGCTTCAAGCTCAACCGGGGCGCAGAGCTGTCCCGGCGCA 218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
650 CGCGGCGCAGTGGAGCTTCAAGCTCAACCGGGGCGCAGAGCTGTCCCGGCGCA 709
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 GAGCTGGGGTTCAGCTGCGAGGGCGCGGCTACCCGGCTCGCGGCTGCGCGC 278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
710 CAGACAGCTTCAAGCTTCCAGAGCTCGGGCTACCGC-----GCTCG 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
279 TCGCGGCGCAGCTGCGGCGCGCGCAGCTGTCTCTCTCTCTCTCTCTCTCTCT 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
752 TCGCGGCGCAGCTGCGGCGCGCGCAGCTGTCTCTCTCTCTCTCTCTCTCTCT 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 CAACTCCAGCGGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
812 CAATTCTACCGGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 CGCGGCGCAGCTGCGGCGCGCGCGCGGCGCGGCT-----GCACACGCGCA 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
872 TAGCGGCGCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 931
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
453 CTCGAGCGCGCAGGCGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
932 TAGGAGCGCGCGGCGCGCGCGCGCGCGCGCTTTTCTGTGGGCTTCAGAGCAG 977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 5
AAK53491 standard; DNA; 114955 BP.

AAK53491:

05-JUL-1999 (first entry)

Human adenosine A1 receptor antisense oligonucleotide fragment.
Antisense oligonucleotide; multiple target; antisense treatment;
impaired respiration; inflammation; lung disease;
pulmonary vasoconstriction; inflammation; allergic rhinitis;
acute asthma; allergy; asthma; impeded respiration;
respiratory distress syndrome; pain; cystic fibrosis;
pulmonary hypertension; pulmonary vasoconstriction; emphysema;
chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
colon cancer; breast cancer; lung cancer; pancreatic cancer;
hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
prostate cancer; ss.

XX Synthetic.
OS
XX W09913886-A1.
PN
XX
XX 25-MAR-1999.
PD
XX
XX 17-SEP-1998; 98WO-US19419.
PF
XX
XX 09-JUN-1998; 98US-0093972.
PR
XX 17-SEP-1997; 97US-0059160.
PX
XX
XX (UYEC-) UNIT EAST CAROLINA.
PA
XX
XX Myce JW.
PI
XX
XX WPI: 1999-229400/19.
DR
XX
XX New antisense oligonucleotides used in treatment of, e.g. pulmonary
PT vasoconstriction
XX
XX
XX Disclosure; Page 37; 120pp; English.

The specification describes antisense oligonucleotides (AAK52869-X55271) directed against at least 2 mRNAs selected from target genes, coding and non-coding regions of mRNAs corresponding to target genes, gene initiation codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-end and the juxta-section between coding and non-coding regions and all segments of mRNAs encoding proteins associated with one or more diseases, conditions or mixtures. The antisense oligonucleotides may be derived from sequences AAK55272-74. These multiple target oligonucleotides (specifically AAK55180-271) can be used for the antisense treatment of diseases and conditions. Typical diseases and conditions are those associated with impaired respiration and inflammation, including lung diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded respiration, respiratory distress syndrome, pain, cystic fibrosis, chronic obstructive pulmonary disease (COPD), and cancers such as leukemia, lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer.

Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other:
Query Match 13.6%; Score 68.2; DB 20; Length 114955;
Best Local Similarity 36.8%; Pred. No. 0.0013;
Matches 175; Conservative 32; Mismatches 267; Indels 1; Gaps 1;

```

15 GGGTCCGCAAGTTCGCGGCGCATCGGGGTGGGAGCTTCGCGCGCGCACCATACACATCCCG 74
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105005 BGGCCBGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105064
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 CGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105065 CGSNNNDNNGCCBGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 GCAGAGCGGCTGCTGTTGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 194
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105125 CGGGCGGNNNDNNGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 CGCGCAGAGCTGCTCCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105185 CGGGCGGNNNDNNGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 CGCGGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105245 GGGCGGNNNDNNGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 TGGCTCTACTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 373
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 105305 NNDNNGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 105364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


XX New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT
 PS Claim 1; Page 100; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
 CC sequence for one such transcription factor. The transcription factor may
 CC be used to produce a plant having modified gene expression such as a
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
 CC mahogany species or to modify the activity of a polypeptide in a plant.
 CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bZIP, bZIP family of G-box
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
 CC and ERBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
 CC and MYB.
 CC
 SQ Sequence 438 BP; 66 A; 196 C; 111 G; 65 T; 0 other;
 Query Match 12.0%; Score 60; DB 21; Length 438;
 Best Local Similarity 51.1%; Pred. No. 0.059;
 Matches 141; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
 DB 39 GGTGGGCACTTCGGCGCCACCAATCACCCTCCGCGCGGCGGCGGCTGCCGC 98
 130 GGGCGGCGCAAGAGCCAGCCAGCAGCAGCCCTCTCCGCGCGGCGGCGGCGG 189
 99 CGAGATGACGAGACCGTGAAGCTGAGCTGGCGGCGGCGAGAGAGCTGCTGATTCCGC 158
 190 TGACGACGACGCGACCACTCGAGACCTCCGCGGCGAGAGACTCTCTCGGCGTCCGG 249
 159 CGCGCGGCGACATGGAGCTTCAAGCTCAACCGGCGGCGGCGAGCTGTCCCGGCGCA 218
 250 GAAGGTCCACGAGTGTGATGTCCCAAGAGCTTCCCAAGCGGCGGCTCGGCGG 309
 DB 219 GAGCTCGGCGTTACGTCGAGGCGCCCGGCGGCGCTCCGCTGCGGCTGCCGC 278
 310 GCACAGCGGCTGCTGATGTCCCAAGAGCTTCCCAAGCGGCGGCTCGGCGG 369
 QY 279 TGGCGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 314
 DB 370 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 405
 RESULT 8
 AAC56714
 ID AAC56714 standard; DNA; 438 BP.
 XX
 AC AAC56714;
 XX
 DT 25-JAN-2001 (first entry)
 XX
 DE Eucalyptus grandis transcription factor DNA sequence #585.
 XX
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
 KW basic helix-loop-helix zipper; homeotic; LIM domain; homeobox; MADS;
 KW homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;
 KW type 2 Cys2His2; CCAAT box element; MYB; ss.
 XX
 OS Eucalyptus grandis.
 XX
 PN WO200053724-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06112.
 XX
 PR 11-MAR-1999; 99US-0266513.

PR 18-AUG-1999; 99US-0149485.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Wood M, McGrath A, Shank MA, Glenn M;
 XX
 DR WPI; 2000-579369/54.
 XX
 PT New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT
 PS Claim 1; Page 498; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
 CC sequence for one such transcription factor. The transcription factor may
 CC be used to produce a plant having modified gene expression such as a
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
 CC mahogany species or to modify the activity of a polypeptide in a plant.
 CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bZIP, bZIP family of G-box
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
 CC and ERBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
 CC and MYB.
 CC
 SQ Sequence 438 BP; 65 A; 195 C; 111 G; 65 T; 2 other;
 Query Match 12.0%; Score 60; DB 21; Length 438;
 Best Local Similarity 51.1%; Pred. No. 0.059;
 Matches 141; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
 DB 39 GGTGGGCACTTCGGCGCCACCAATCACCCTCCGCGCGGCGGCGGCTGCCGC 98
 130 GGGCGGCGCAAGAGCCAGCCAGCAGCAGCCCTCTCCGCGGCGGCGGCGGCGG 189
 99 CGAGATGACGAGACCGTGAAGCTGAGCTGGCGGCGGCGGAGAGAGCTGCTGATTCCGC 158
 190 TGACGACGACGCGACCACTCGAGACCTCCGCGGCGAGAGACTCTCTCGGCGTCCGG 249
 159 CGCGCGGCGACATGGAGCTTCAAGCTCAACCGGCGGCGGCGAGCTGTCCCGGCGCA 218
 250 GAAGGTCCACGAGTGTGATGTCCCAAGAGCTTCCCAAGCGGCGGCTCGGCGG 309
 DB 219 GAGCTCGGCGTTACGTCGAGGCGCCCGGCGGCGCTCCGCTGCGGCTGCCGC 278
 310 GCACAGCGGCTGCTGATGTCCCAAGAGCTTCCCAAGCGGCGGCTCGGCGG 369
 QY 279 TGGCGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 314
 DB 370 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 405
 RESULT 9
 AAL36353/C
 ID AAL36353 standard; DNA; 9968 BP.
 XX
 AC AAL36353;
 XX
 DT 08-JAN-2002 (first entry)
 XX
 DE Human musculoskeletal system related polynucleotide SEQ ID NO 2718.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system; ds.

OS Homo sapiens.
 XX W0200155367-A1.
 XX
 PD 02-AUG-2001.
 PF 17-JAN-2001; 2001WO-US01338.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225271.
 PR 14-AUG-2000; 2000US-0225275.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227189.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0232968.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0234300.
 PR 14-SEP-2000; 2000US-0234301.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR

PR 27-SEP-2000; 2000US-0235634.
 PR 27-SEP-2000; 2000US-0235636.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

PA (DZIE/) DZIEGLEMSKA H.
 PA (ZOTC/) ZOTCHEV S B.
 PA (SEKU/) SEKUROVA O N.
 PA (FJAE/) FJAEVRIK E.
 PA (BRAU/) BRAUTASET T.
 PA (STRO/) STROM A R.
 XX Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
 PI Valla S, Ellingsen TE, Sletta H, Gulliksen O;
 XX
 DR WPI: 2001-557614/62.
 DR P-PSDB: AAE10125, AAE10126, AAE10127, AAE10128, AAE10129, AAE10130,
 DR AAE10131, AAE10132, AAE10133, AAE10134, AAE10135, AAE10136, AAE10137.
 XX
 PT New nystatin polyketide synthase polynucleotides and polypeptides,
 PT useful as antibiotics and antifungals -
 XX
 PS Claim 2; Page 116-151; 266pp; English.
 XX
 CC The present invention relates to the cloning and sequencing of the gene
 CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
 CC involved in the biosynthesis of the macrocyclic antibiotic nystatin.
 CC The nystatin PKS is useful as antifungal antibiotics. The present
 CC sequence is a Streptomyces noursei nysI DNA of nystatin PKS gene cluster.
 XX
 SQ Sequence 65140 BP; 8270 A; 25171 C; 22273 G; 9426 T; 0 other;
 XX
 Query Match 11.3%; Score 56.6; DB 22; Length 65140;
 Best Local Similarity 44.6%; Pred. No. 0.12;
 Matches 221; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
 XX
 QY 2 TGCCTCCGAGCGGGGTCGCGACGTTCCCGCATGCGGGGTGGCGACGTTCCGCGCCACG 61
 DB 18987 TGCCTCCGAGCGGGGTCGCGACGTTCCCGCATGCGGGGTGGCGACGTTCCGCGCCACG 61
 QY 62 ATCAACACTCCG 121
 DB 19047 GCAAGGACCG 19106
 QY 122 GCCCGGCG 181
 DB 19107 GCGTCACG 19166
 QY 182 AGCTCAACCG 241
 DB 19167 CCACTACCG 19226
 QY 242 GCCCGGCG 301
 DB 19227 TCACCG 19286
 QY 302 ACG 361
 DB 19287 CCGACG 19346
 QY 362 TCCGAGCG 421
 DB 19347 CCGACG 19406
 QY 422 GCG 481
 DB 19407 TCCCG 19466
 QY 482 CGGCGCTGCCAGAC 496
 DB 19467 TGGTGTGCGCCGAGC 19481
 XX
 RESULT 11
 ID AAD17186 standard; DNA; 125401 BP.
 XX
 AC AAD17186;

XX
 DT 29-NOV-2001 (first entry)
 XX
 DE Streptomyces noursei nystatin PKS gene cluster DNA.
 XX
 KW polyketide synthase; PKS; macrocyclic; nystatin; PKS gene cluster;
 KW antifungal; antibiotic; ds.
 XX
 OS Streptomyces noursei.
 XX
 FH Key Location/Qualifiers
 FT CDS 6337..34771
 FT /tag= a
 FT /product= "NysI complete protein"
 FT 34792..51099
 FT /tag= b
 FT /product= "NysJ protein"
 FT 51155..57355
 FT /tag= c
 FT /product= "NysK protein"
 FT 57503..58687
 FT /tag= d
 FT /product= "NysL protein"
 FT complement (58786..58980)
 FT CDS
 FT /tag= e
 FT /product= "NysM protein"
 FT /note= "CDS does not include start codon"
 FT complement (59045..60241)
 FT /tag= f
 FT /product= "NysN protein"
 FT /note= "CDS does not include start codon"
 FT complement (60238..61296)
 FT CDS
 FT /tag= g
 FT /product= "NysD2 complete protein"
 FT 120628..121308
 FT /tag= h
 FT /product= "NysR4 (long) protein"
 XX
 PN WO200159126-A2.
 PD 16-AUG-2001.
 XX
 PF 08-FEB-2001; 2001WO-GB00509.
 XX
 PR 08-FEB-2000; 2000GB-0002840.
 PR 10-APR-2000; 2000GB-0008786.
 PR 14-APR-2000; 2000GB-0009387.
 XX
 PA (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPSELIGE.
 PA (SNTF) SINTEF STRIPELSEN IND TEK FORSK.
 PA (ALPH-) ALPHARMA AS.
 PA (STNV-) STINVENT AS.
 PA (DZIE/) DZIEGLEMSKA H.
 PA (ZOTC/) ZOTCHEV S B.
 PA (SEKU/) SEKUROVA O N.
 PA (FJAE/) FJAEVRIK E.
 PA (BRAU/) BRAUTASET T.
 PA (STRO/) STROM A R.
 XX
 PI Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
 PI Valla S, Ellingsen TE, Sletta H, Gulliksen O;
 XX
 DR WPI: 2001-557614/62.
 DR P-PSDB: AAE10143, AAE10144, AAE10145, AAE10146, AAE10147, AAE10148,
 DR AAE10149, AAE10150.
 XX
 PT New nystatin polyketide synthase polynucleotides and polypeptides,
 PT useful as antibiotics and antifungals -
 XX
 PS Claim 1; Page 188-254; 266pp; English.
 CC The present invention relates to the cloning and sequencing of the gene
 CC cluster encoding a modular type I polyketide synthase (PKS) enzyme

CC Involved in the biosynthesis of the macrolide antibiotic nystatin.
 CC The nystatin PKS is useful as antifungal antibiotics. The present
 CC sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.
 XX

SQ Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other;

Query Match 11.3%; Score 56.6; DB 22; Length 125401;
 Best Local Similarity 44.6%; Pred. No. 0.11;
 Matches 221; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

2 TGGTCTGAGACGCGGTCGACGATCCCGCCATCGGGGTGGGACGTTTCGCGCCACG 61
 79248 TGGTCTGAGACGCGGTCGACGATCCCGCCATCGGGGTGGGACGTTTCGCGCCACG 79307
 62 ATACACCACTCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
 79308 GCAAGACG 79367
 122 GCGTGGCG 181
 79368 GCGTGGCG 79427
 182 AGGTCAACCGCGGCG 241
 79428 CCACTTACCG 79487
 242 GCG 301
 79488 TCACCG 79547
 302 ACAGCG 361
 79548 CCGAGCG 79607
 362 TCGCGACG 421
 79608 CCGAGCG 79667
 422 GCG 481
 79668 TCG 79727
 482 CGGCG 496
 79728 TGGTGGCG 79742

RESULT 12
 AAA09686 standard; DNA; 3957 BP.
 XX
 AC AAA09686;
 XX
 DT 31-JAN-2001 (first entry)
 XX
 DE HSV-2 immediate early protein ICP4 DNA sequence.
 XX
 KW Herpes-simplex-virus type 2; HSV-2; infected cell protein 4; ICP4;
 KW vaccine; infection; ds.
 OS Herpes simplex virus type 2.
 XX
 PN WO9516779-A1.
 XX
 PD 22-JUN-1995.
 XX
 PF 13-DEC-1994; 94MO-BP04138.
 XX
 PR 14-DEC-1993; 93GB-0025496.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Pala P, Gheysen DR, Slaoui MM, Koutoukos MC;

XX WPI: 2001-024142/03.
 DR P-PSDB; AAB26874.
 XX

PT Immediate early herpes-simplex-virus type 2 (HSV-2) ICP4 protein is
 PT used in vaccines for therapeutically or prophylactically treating HSV
 PT infections -
 XX
 PS Claim 5; Page 16; 28pp; English.
 XX

CC This invention relates to an immediate early herpes-simplex-virus type 2
 CC (HSV-2) infected cell protein 4 (ICP4) recognised by human cytotoxic T
 CC cells. HSV-2 ICP4 protein is recognized by cytotoxic T-lymphocyte (CTL)
 CC cells in humans and is used in vaccines for therapeutically or
 CC prophylactically treating HSV infections. Pharmaceutical compositions of
 CC HSV-2 ICP4 protein may be used to treat patients suffering from HSV
 CC infections, to prevent or decrease recurrent herpes disease, frequency,
 CC severity and duration of episodes. The present sequence represents HSV-2
 CC DNA encoding ICP4.
 XX

SQ Sequence 3957 BP; 368 A; 1566 C; 1568 G; 365 T; 0 other;

Query Match 11.3%; Score 56.4; DB 22; Length 3957;
 Best Local Similarity 45.6%; Pred. No. 0.18;
 Matches 198; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

54 GCGCACCAATCACACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 113
 2127 GCGCTCG 2186
 114 TGAACG 173
 2187 GCGCGGACG 2246
 174 AGCTTCAAGCTTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 233
 2247 GACCTTCCG 2306
 234 GTGCGAGGCG 2393
 2307 GCGGCG 2366
 294 CGGCG 353
 2367 GCG 2426
 354 CTTCCTGTTCCGAGCG 413
 2427 CCGGAAGAGAGCG 2486
 414 GCTGTTGCG 473
 2487 GCG 2546
 474 CCGTTCG 487
 2547 GCGCGCGCGAGGCGC 2560

RESULT 13
 AAD25519 standard; DNA; 154746 BP.
 XX
 AC AAD25519;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human herpesvirus 2 complete DNA genome.
 XX
 KW Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;
 KW antibacterial; fungicide; protozoicide; antipneumatic; antiinflammatory;
 KW antitubercular; rheumatoid arthritis; neuroproliferative; multiple sclerosis;
 KW immune response; vasotrophic; vaccine; gene therapy; autoimmune disease;

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 1667 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Y-chromosome, SRY gene
US-08-196-016A-2

Query Match 26.4%; Score 132; DB 1; Length 1667;
Best Local Similarity 71.3%; Pred. No. 3.8e-16;

Matches 174; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 201 CCAGACCGGGTCAAGAGCCCATGAAAGCCTTCATGCTGTGTGTCCTCCGCGGACGCGC 260
DB 792 CCAGATGCTGTCAGCAGCCCATGAAAGCCTTCATGCTGTGTGTCCTCCGATCAAGAGAG 851
QY 261 CAAGATGCGCCAGAGAACCCCAAGATGACACACTCGAGATCAAGCAGCCTGGGGGC 320
DB 852 AAAAGTGGCTTAGAGAACCCCAAGATGACACACTCGAGATCAAGCAGCCTGGGGATG 911
QY 321 CGAGTGGAGTCAATGTCGAGGCGGAGAGAGCGGCTTCATCGACGAGGCGCAAGCGCT 380
DB 912 CAAGTGGAAATGCTTACAGAAAGCCGAAAGCGCCCATTCCTTCAGGAGGCGACAGAGCT 971
QY 381 GCGGCGCTGTCATGAGAGGACCCCGATTCAGATACCGGCGCGCCCAAGACCA 440
DB 972 ACAGCGGCTGACGAGATTAATACCGGCGCTATTAATACCGAGCTCTGCGCAAGGAGA 1031
QY 441 GAGC 444
DB 1032 GAGG 1035

RESULT 6

US-08-196-016A-1

Sequence 1, Application US/08196016A

Patent No. 5596089

GENERAL INFORMATION:

APPLICANT: SILVERSIDES, David W.

APPLICANT: DANEAU, Isabelle M.F.

APPLICANT: HOUE, Alain

TITLE OF INVENTION: SRY GENOMIC SEQUENCES AND USE IN TISSUE

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: Continental Plaza, 411 Hackensack Avenue

CITY: Hackensack

STATE: N.J.

COUNTRY: U.S.A.

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/196,016A

FILING DATE: 14-FEB-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: JACKSON, David A.

REGISTRATION NUMBER: 26,742

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Y-chromosome, SRY gene
US-08-196-016A-1

Query Match 24.3%; Score 124.6; DB 1; Length 1661;
Best Local Similarity 67.6%; Pred. No. 8e-15;

Matches 175; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 194 AGGCAACCAAGAGAGCGGCTCAAGAGCCCATGAAAGCCTTCATGAGTGTGTCCTCCGCGG 253
DB 1054 AGAGCAGCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1113
QY 254 AGCGGCGCAAGATGCGCCAGAGAGAACCCCAAGATGCAACAATCGAGATGAGCAAGCGC 313
DB 1114 GAAGACGAAAGAGTGTCTTAGAGATCCCAAAATGAAATGAAATGCAATCAGACATGAGAGAGC 1173
QY 314 TGGGGCGCGAGTGAAGTCAATGTCGAGGCGGAGAGAGCGGCGGCTTCATGAGAGAGGCA 373
DB 1174 TGGATATGATGAGTGAAGTCAATGTCGAGGCGGAGAGAGCGGCGGCTTCATGAGAGAGGCA 1233
QY 374 AGCGGCTGCGCGCGCTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
DB 1234 AGAGACTACTAGATCAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1293
QY 434 AGACCAAGAGCGCTGCTCA 452
DB 1294 GAGCCAGAGAGCGCACAGAA 1312

RESULT 7

US-08-239-276-6

Sequence 6, Application US/08239276

Patent No. 5840836

GENERAL INFORMATION:

APPLICANT: Rabin, Daniel

TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sprung Kramer Schaefer & Briscoe

STREET: 660 White Plains Road

CITY: Tarrytown

STATE: New York

COUNTRY: USA

ZIP: 10591-5144

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MD storage

COMPUTER: Apple Macintosh

OPERATING SYSTEM: System 7.5

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/239,276

FILING DATE: 05-MAY-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/872,646

FILING DATE: 08-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/715,181

FILING DATE: 14-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/441,703

FILING DATE: 04-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/312,543

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

QY	82	AACGACGTCACATCGCCCGCGGGGCGCCAGGCCCGCCAGAACCTCTCGGCGCCCGCGGG	141
Db	890	ATCACCTCGACACTCATCCCGCCAGGAGCGGCTGTGAGAGACGGCTGTGTGCACCCACTG	949
QY	142	GCGGCGCGCGCGCGGGGGGAGGCGGGGGCGCGCGCGCGCGGGCGCCAAAGGCCAAC	201
Db	950	GAGGAGGCCATGCTGAGCTGTGGCATGGATGGCTCCGCCACTTCCCGAGTCCCGAAAC	1000
QY	202	CAGGACCGGGTCAAAAGCGCCCATGACGGCTTTCATGTGTGTGTGTCCCGGAGACGGCGC	261
Db	1010	AGCACCCATCATCAAGAGGCCCATGACGCCCTTTCATGTGTGTGTGTGCGCAAGATGAGCGAGG	1060
QY	262	AGATGTGGCCCGAGGAGAACCCCAAGTGTGCACACCTCGGAGATGCACAGACGGCGTGGGGGCC	321
Db	1070	AGATCTCTGCAAGCCTTCCAGACATGTGCACACTCTCAGCATCAGCAAGATCTTTGATCT	112
QY	322	GAGTGGAGGTCTATCTCCGAGGCCGAGAAAGGGCGCTTTCATGCACAGAGGCCAAGCGCTG	381
Db	1130	CGCTGGAATGTCATATACCAACACAGGAGAGAGAGCCCTACTATGAGAGAACAGCGCGCTG	118
QY	382	CGCGCGCTGCACATGAGGAGACCCCGGATTTACAAGTACCGCGCCGCGCGCAAGACCA	439
Db	1190	ACCCGGCAGACCTGTGAGAAATATCTGTACTTACACAGTACAGACGCCCGCGCCCAAGCGCA	1247

RESULT 9
 US-08-468-577B-6
 ; Sequence 6, Application us/08468577B
 ; Patent No. 6001804
 ; GENERAL INFORMATION:
 ; APPLICANT: Rabin, Daniel
 ; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENESIS
 ; TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Spring Kramer Schaefer & Briscoe
 ; STREET: 660 White Plains Road
 ; CITY: Tarrytown
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10591-5144
 ; COMPUTER READABLE FORM:

LENGTH: 1397 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-468-579B-1

Query Match 17.0%; Score 84.8; DB 2; Length 1397;
Best Local Similarity 65.3%; Pred. No. 1.1e-07;
Matches 139; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

QY 218 GGCACATGAAGCCCTTCATGCTGTCCTCCGCGGCGCAAGATGCCCCAGAGA 277
DB 1 GGCACATGAAGCCCTTCATGCTGTCCTCCGCGGCGCAAGATGCCCCAGAGA 60
QY 278 ACCCAAGATGCACAACTCGAGATCAGCAGCGCTGGGGCCGAGTGGAGTCTATGT 337
DB 61 TCCAGACATGCACAACTCCAGCATCAGCAGAGATCTTGATCTCGCTGGAATCCATGA 120
QY 338 CCGAGCGCGAGAGCGGCGCTTCATCGACAGGCCAAGCGGCTGGCGCGCTGCACATGA 397
DB 121 CCAACAGAGAGAGAGCGGCGCTTCATCGAGAGAGAGCGCTGCTGATCG-TCAATCAGCTGG 179
QY 398 AGAGACACCGGATTCATCAAGTACCGCGCGCGCC 430
DB 180 AGAGTATCTGACTACATCAAGTACAGCGCGCGC 212

RESULT 15
US-08-468-577B-1
Sequence 1, Application US/08468577B
Patent No. 6001804

GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,577B
FILING DATE: 06-JUN-1995

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kurt G. Briscoe

REGISTRATION NUMBER: 33,141

REFERENCE/DOCKET NUMBER: MDI 251.8-KGB

TELECOMMUNICATION INFORMATION:

TELEPHONE: (914) 332-1700

TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1397 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-468-577B-1

Query Match 17.0%; Score 84.8; DB 3; Length 1397;
Best Local Similarity 65.3%; Pred. No. 1.1e-07;
Matches 139; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

QY 218 GGCACATGAAGCCCTTCATGCTGTCCTCCGCGGCGCAAGATGCCCCAGAGA 277
DB 1 GGCACATGAAGCCCTTCATGCTGTCCTCCGCGGCGCAAGATGCCCCAGAGA 60
QY 278 ACCCAAGATGCACAACTCGAGATCAGCAGCGCTGGGGCCGAGTGGAGTCTATGT 337
DB 61 TCCAGACATGCACAACTCCAGCATCAGCAGAGATCTTGATCTCGCTGGAATCCATGA 120
QY 338 CCGAGCGCGAGAGCGGCGCTTCATCGACAGGCCAAGCGGCTGGCGCGCTGCACATGA 397
DB 121 CCAACAGAGAGAGAGCGGCGCTTCATCGAGAGAGAGCGCTGCTGATCG-TCAATCAGCTGG 179
QY 398 AGAGACACCGGATTCATCAAGTACCGCGCGCGCC 430
DB 180 AGAGTATCTGACTACATCAAGTACAGCGCGCGC 212

Search completed: October 10, 2002, 17:33:14
Job time : 31.7333 secs


```
/clone_lib="UT-R-CA0"
/lab_host="DH10B (Life Technologies)"
/Note="Vector: pTR73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I. Site 2: Eco RI. The UT-R-CA0
library is a subdirected library derived from the following
tissues: thalamus, cerebellum, hypothalamus, medulla, pons
, midbrain, cerebral cortex, corpus striatum, testis, and
hippocampus. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratat.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
genome research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT      51 a      186 c      144 g      86 t
ORIGIN

Query Match      56.2%; Score 280.8; DB 9; Length 467;
Best Local Similarity 96.0%; Pred. No. 1.6e-28;
Matches 288; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 201 CCAGGACGGGTCAAAACGGCCCATGAAAGCCTTCATGTGTGTCGGGGGAGCGGGG 260
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 CCAGATGAGGGGTCAAGCGCCCATGAAAGCCTTCATGTGTGTCGGGGGAGCGGGG 408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 261 CAGATGGCCCAAGGAAACCCCAAGATGCAACATCGGAGATCAGCAAGCGCTGGGGG 320
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 CAGATGGCCCAAGGAAACCCCAAGATGCAACATCGGAGATCAGCAAGCGCTGGGGG 348
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 321 CGAGTGAAGGATCATGTGCGAGGCGGAGAAAGCGCGTCATCGAGAGGCGCAAGCGCT 380
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 CGAGTGAAGGATCATGTGCGAGGCGGAGAAAGCGCGTCATCGAGAGGCGCAAGACT 288
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 381 GCGCGGCTGCACATGAAGAGACCCCGGATTACAGTACCGCGCCGCGCAAGACCA 440
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 GCGCGGCTGCACATGAAGAGACCCCGGATTACAGTACCGCGCCGCGCAAGACCA 228
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 441 GACGCTGTCAGAAAGGACAACTACTCGTGGCGCGCGGCTCTTGGCGCGCGCGG 500
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 GACGCTGTCAGAAAGGACAACTACTCGTGGCGCGCGGCTCTTGGCGCGCGG 168
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
BF305922      739 bp      mRNA      linear      EST 21-NOV-2000
LOCUS      601889206F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4123153 5',
DEFINITION      mRNA sequence.
ACCESSION      BF305922
VERSION      BF305922.1 GI:11252955
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 739)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgaps-remail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Ling Hong/Rubin Laboratory
      CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.ni.gov
      Plate: LLCM1007 row: 1 column: 02
      High quality sequence stop: 677.
      Location/Qualifiers
      1..739
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:4123153"
      /clone_lib="NIH_MGC_17"
```

```
/tissue_type="rhadomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: muscle; Vector: pORF1; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(C) size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
BASE COUNT      160 a      231 c      249 g      99 t
ORIGIN

Query Match      53.1%; Score 265.6; DB 10; Length 739;
Best Local Similarity 84.7%; Pred. No. 1.6e-26;
Matches 298; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 145 GCGGCGCGCGGGGCGGAGCGGGGCGGGCGGGCGGGGCGCCCAAGGCCCAACCG 204
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 GGGGGGCGGGCGGGCGCAACTCCACCGCGCGCGCGCGGCAACCAAGAAACACCGCG 91
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 205 GACCGGGTCAAAAGCGCCCATGAAAGCCTTCATGTGTGTCGGGGGAGCGGGCGCAAG 264
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 GACCGGCTCAAGCGGGCCATGATGCTTTCATGTGTGTCGGGGGAGCGGGCGCAAG 151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 265 ATGGCCCAAGAAACCCCAAGATGCAACATCGGAGATCAGCAAGCGCTGGGGCGGAG 324
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 ATGGCCCAAGAAACCCCAAGATGCAACATCGGAGATCAGCAAGCGCTGGGGCGGAG 211
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 325 TGAAGGTGATGTGCGAGGCGGAGAAAGCGCGTCATCGAGAGGCGCAAGCGCGCTGCC 384
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 TGAAGGTGATGTGCGAGGCGGAGAAAGCGCGTCATCGAGAGGCGCGCTGCCA 271
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 385 GCGGTGCACATGAAGAGACCCCGGATTACAGTACCGCGCGCGCGCAAGACCAAGAG 444
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 GCGGTGCACATGAAGAGACCCCGGATTACAGTACCGCGCGCGCGCAAGACCAAGAG 331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 445 CTGCTCAAGAGACAACTACTCGTGGCGCGCGGCTCTTGGCGCGCGGCGG 496
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 CTCATGAAGAGATAGTACACCGCTGCCGCGGCTCTTGGCGCGCGGCGG 383
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
A1356682      965 bp      mRNA      linear      EST 15-FEB-1999
LOCUS      AY228210.x1 NCI-CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2012730 3',
DEFINITION      similar to gb:U71136 SOX-10 PROTEIN (HUMAN);, mRNA sequence.
ACCESSION      A1356682
VERSION      A1356682.1 GI:4108303
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 965)
AUTHORS      NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES      National Cancer Institute / National Institute of Neurological
JOURNAL      Disorders and Stroke, Brain Tumor Genome Anatomy Project
COMMENT      Unpublished (1998)
      Contact: Robert Strausberg, Ph.D.
      Email: cgaps-remail.nih.gov
      Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
      Ph.D.
      CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
      Bonaldo, Ph.D.
      CDNA Library Arrayed by: Greg Lennon, Ph.D.
      DNA Sequencing by: Washington University Genome Sequencing Center
      Clone distribution: NCI-CGAP clone distribution information can be
      found through the I.M.A.G.E. Consortium/LNLN at:
      www.bio.lnl.ni.gov/bbrp/image/image.html
      Insert length: 1190 Std Error: 0.00
      Seq primer: -40UP from Gibco
```


[illegible]

Db	366	CGCTCATGACAGAGATTAAGTACACGCTGCCGGCGGCCTGCTGCGCCCCGAGC	419
RESULT#	8		
LOCUS	AM005368		
DEFINITION	AM005368	460 bp	mRNA
ACCESSION	wz55h12.x1	NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:256571.3'	linear EST 20-Oct-2000
VERSION	similar to gb:X71136 SOX-10 PROTEIN (HUMAN);		mRNA sequence.
KEYWORDS	AM005368.1	GI:5854146	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 460)		
COMMENT	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BIGAP), Tumor Gene Index Unpublished (1998) Contact: Robert Strausberg, Ph.D. Email: cgapbs-ftemail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/dbfp/image/image.html Insert length: 1157 Std Error: 0.00 Seq primer: -40mp from Gibco High quality sequence stop: 457. Location/Qualifiers 1..460 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="IMAGE:256571.3" /clone_lib="NCI_CGAP_Brn25" /tissue_type="anaplastic oligodendroglioma" /lab_host="TDH10B" /note="Organ: brain; Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTCCACTGTGAAGTGCGAGCGCGCCGCAAGATGTTTTTTTTTTTTTTT 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ19 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."		
FEATURES	source		
BASE COUNT	104 a	140 c	165 g
ORIGIN			51 t
Query Match	52.9%;	Score 264.4;	DB 9; Length 460;
Best Local Similarity	84.2%;	Pred. No. 2.3e-26;	
Matches 298;	Conservative	0;	Mismatches 56; Indels 0; Gaps 0;
OY	143	CGGGGGGGGGGGGGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGAGGCCAACCC	202
Db	66	CGGGGGGGGGGGGGGCACTCACCGCGGGCGGGCGGGCAACGAAAAAAGAGCC	125
OY	203	AGGACCGGGTCAAACGGCCCATGAACGGCTTCATGGTGTGTGCCGGGCGAGCGGCGCA	262
Db	126	CGGACCGGCTCAAGGGGCCCATGAAATGCTTCATGGTGTGTGCCGGGCGAGCGGCGCA	185
OY	263	AGATTGGCCCGAGAGAACCACAAGATGCAACAACCTGGAGTGTGCAAGAGCGCTCGGGGGCGCG	322
Db	186	AGATTGGCCCGAGAGAACCACAAGATGCAACAACCTGGAGTGTGCAAGAGCGCTCGGGGGCGCG	245
OY	323	ACTGGAAGTCAATGTCGAGGCGGAGAGAGCGGCGCTTCATGCAGCAGGCGCAAGCGGCTGC	382

RESULT 14	
LOCUS	A1421743 483 bp mRNA linear EST 30-MAR-1999
DEFINITION	tcf9pb01.x1 NCI-CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2098537 3' similar to gp.X71136 SOX-10 PROTEIN (HUMAN); contains MER22.L1 TARI repetitive element ; , mRNA sequence.
ACCESSION	A1421743
VERSION	A1421743.1 GI:4267674
KEYWORDS	EST.
SOURCE	Homo sapiens human.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 483) NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap/ . National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAIP), Tumor Gene Index Unpublished (1998)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
COMMENT	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at: www.bio.linnl.gov/bdtp/Image/Image.html Insert Length: 1154 Std Error: 0.00 Seq primer: -400P from Gibco High quality sequence stop: 480. Location/Qualifiers 1..483 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2098537" /clone_lib="NCI-CGAP_Brn23" /lisse_type="gliblastoma (pooled)" /_lab_host="DH10B" /note="Organ: brain; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site1: Not I - oligo(dT) primer [5' TGTTACCATTCTGAAGTGGGGCGGCCGCAGATCTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	104 a 152 c 175 g 51 t 1 others
ORIGIN	
Query Match	52.9% Score 264.4; DB 9; Length 483;
Best Local Similarity	84.2%; Fred. No. 2.3e-26;
Matches 298:	Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY	143 CGGGCGGGGGCGGGGCGGAGCGGGGGCGGGCGGGCGGGGGCGCCCAAGCACC 202
DB	91 CGGGGGGGGGCGGGCGCAACTCACCGGGGGCGGGCGGGCGCAACGAAAAACAAGC 150
QY	203 AGGACCGGGTCAAACGGCGCCATGAACCGCTTCATGTGTGTGCCGGGGCGCGCA 262
DB	151 CGGACCGGCTCAAGCGGCCCATGAATGCTTCATGTGTGTGCCGGCGGCAAGCGCA 210
QY	263 AGATGCGCCAGGAGAAGCCCAAGATGCAACAATCGAGATGAGCAAGCGCTGGGGCGC 322
DB	211 AGATGCGCCAGGAGAAGCCCAAGATGCAACAATCGAGATGAGCAAGCGCTGGGGCGC 270
QY	323 AATGGAAGTCAATGTCGAGGCGGAGAACGGCGCTTCATGCAGAGGCGCAAGCGCTGC 382

QY	DB	271	AGTGAACCTTTTGGTCGGAGACGGAGACGGCGCCCTTCATCGACAGACGCTAAGCGCTGC	330
QY	383	GCAGCCTGCACATGAAAGAGACACCCGGATTCACGATACCAGCGCGCGCGCAAGACCAAGA	442	
Db	331	GAGCGCTCCACATGTAAGAGACACCCGGATTTAAATTAATCCCGCCCGCGGAAACCAAGA	390	
QY	443	CGCTGCTCAGAGAGACAACTACTGCTGGCGCGCGCGGCTCTGCGCGCGCGCGC	496	
Db	391	CGCTCATGAGAGAGATTAAGTACACGCTGCGCGCGCGGCTCTGCGCGCGCGCGC	444	
RESULT 15				
LOCUS	AM028031			
DEFINITION	AM028031	485 bp	mrna	linear
ACCESSION	AM028031			EST 08-MAR-2000
VERSION	AM028031.1	GI:5886787		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	human.			
REFERENCE	1 (bases 1 to 485)			
AUTHORS	NCBI/NIHDS-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTCAP), Tumor Gene Index			
JOURNAL	Unpublished. (1998)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: www-bio.illnl.gov/dbfp/image/image.html Insert Length: 1189 Std Error: 0.00 Seq primer: -400P from Glpco High quality sequence stop: 456. Location/Qualifiers 1. 485 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2501832" /clone_id="NCI-CGAP_Brn23" /issue_type="gliblastoma (pooled)" /lab_host="DH10B" /note="Organ: brain; Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I - Site:2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGGTACCAATCTGAAGTGGAGCGCGCGCATCTTTTATTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	109 a	145 c	176 g	54 t 1 others
ORIGIN				
Query Match	52.9%	Score 264.4	DB 9	Length 485
Best Local Similarity	84.2%	Prod. No. 2.3e-26		
Matches 298	Conservatively	0	Mismatches 56	Indels 0
			Gaps	0
QY	143	CGAGCGCGCGCGCGCGAGAGCGGAGCGCGCGCGCGCGCGCGCGCGCAAGCCCAAC	202	
Db	45	CGAGCGCGCGCGCGCGCACTCACCGCGCGCGCGCGCGCGCGCAACAGAAAAACAC	104	
QY	203	AGGACCGGAGTCAACGCGCCATGAAGCGCTTCATGATGTGTGTCGCCGCGGACGCGGCGA	262	

```

Db      105  CGGACCGCGTCAGAGCGGCCCATGATGCTTCATGTTGTGTGCCCGGGGACAGCGGCGCA 164
QY      263  AGATGGCCCCAGAGAACCCCAAGATGCACAACTCGGAGATCAGCAAGCGCTGGGGGCGG 322
Db      165  AGATGGCCCCAGAGAACCCCAAGATGCACAACTCGGAGATCAGCAAGCGCTGGGGGCGG 224
QY      323  AGTGAAGGTTCATGTCCGAGGCGGAGAGGCGGTCATCGAGAGGCCCAAGCGGCTGC 382
Db      225  AGTGAAGGTTCATGTCCGAGGCGGAGAGGCGGTCATCGAGAGGCCCAAGCGGCTGC 284
QY      383  GCGCGCTGCACATGAGAGGACCCCGGATTACAACTACCGCGCGCGCGGCAAGCAAGA 442
Db      285  GAGCGCTGCACATGAGAGGACCCCGGATTACAACTACCGCGCGCGCGGCAAGCAAGA 344
QY      443  CGCTGCTCAAGAAAGACAAGTACTGCGCGCGCGGCTCTGCGCGCGCGGCG 496
Db      345  CGCTCATGAAGAAAGATTAAGTACGCTGCGCGCGGCTCTGCGCGCGCGGCG 398

```

Search completed: October 10, 2002, 20:27:16
 job time : 698.733 secs

XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
 XX WPI; 2001-457597/49.
 DR P-PSDB; AA05811.
 XX
 PT Isolated polypeptide, used to treat or prognose a disorder
 PT characterized by expression of a hCAP e.g. cancer, is encoded by an
 PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
 PS Claim 57; Page 88-90; 152pp; English.
 XX
 CC The invention relates to nucleic acids and encoded polypeptides which
 CC are cancer associated antigens expressed in patients afflicted with
 CC small cell lung cancer. The molecules provided by the invention can be
 CC used in the diagnosis, monitoring, research or treatment of conditions
 CC characterized by the expression of one or more cancer associated
 CC antigens. The polypeptide is used to treat a disorder characterized by
 CC expression of a hCAP, and determine regression, progression or onset
 CC of a condition characterized by expression of an abnormal amount of a
 CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
 CC are small and non-small cell lung cancer, melanoma, colon, breast, head
 CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
 CC The present sequence is a small cell lung cancer associated gene
 CC designated as NY-SCIC-2, encoding human SOX1 protein.
 XX
 SQ Sequence 4091 BP; 924 A; 1109 C; 1110 G; 943 T; 5 other;
 Query Match 100.0%; Score 500; DB 22; Length 4091;
 Best Local Similarity 100.0%; Pred. No. 9.6e-70;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGGCGCTATGCTCCAGGCGCTCTCTCGGCGGTGATACCCGCGACCGCGCG 60
 DB 1 CCGGCGCTATGCTCCAGGCGCTCTCTCGGCGGTGATACCCGCGACCGCGCG 60
 QY 61 ATGTACAGCATGATGATGAGACCGACTCTCGCGCGCGCGCGCGCGCGCG 120
 DB 61 ATGTACAGCATGATGATGAGACCGACTCTCGCGCGCGCGCGCGCGCGCG 120
 QY 121 AACCTCGGCG 180
 DB 121 AACCTCGGCG 180
 QY 181 GCGGCGGCG 240
 DB 181 GCGGCGGCG 240
 QY 241 TGGTCCCG 300
 DB 241 TGGTCCCG 300
 QY 301 ATAGCAGAGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 DB 301 ATAGCAGAGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 ATGAGCAGAGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 DB 361 ATGAGCAGAGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 QY 421 CCG 480
 DB 421 CCG 480
 QY 481 CTCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500
 DB 481 CTCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500

RESULT 2
 AAX16151
 ID AAX16151 standard; cDNA to mRNA; 1542 BP.
 XX

AC AAX16151;
 XX
 DT 22-APR-1999 (first entry)
 XX
 DE Human Sox1 encoding cDNA.
 XX
 KW Human; Sox1: neuronal stem cell gene; neuroblastic cell; cancer;
 KW nervous system; neurological disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT misc_difference 1..60 /tag- a
 FT /note= "specifically claimed Sox1 control sequence"
 FT CDS 60..1223 /tag- b
 FT /product= "Sox1"
 XX
 PN W09900516-A2.
 XX
 PD 07-JAN-1999.
 XX
 PF 25-JUN-1998; 98MO-GB01862.
 XX
 FR 25-JUN-1997; 97GB-0013469.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 XX
 PI Lovell-Badge R, Pevery LH, Smith A;
 XX
 DR WPI; 1999-095759/08.
 XX
 DR P-PSDB; AAM94461.
 XX
 PT Isolating neuroblastic cells from a population - by detecting the
 PT expression of the Sox1 gene in the cells and sorting the cells to
 PT isolate those cells expressing Sox1
 XX
 PS Claim 13; Page 45-47; 60pp; English.
 XX
 CC A method has been developed for isolating neuroblastic cells from a cell
 CC population. The method comprises: (a) detecting Sox1 gene expression in
 CC the cells; and (b) isolating those cells expressing Sox1. Also described
 CC is a method for producing a cell committed to the neuronal lineage,
 CC comprising: (a) transfecting a pluripotent stem cell with a genetic
 CC construct encoding Sox1 expression; (b) culturing the stem cells to
 CC differentiate into neural cells; and (c) isolating those neural cells
 CC produced. The present sequence encodes human Sox1. Detection of Sox1
 CC expressing cells is important in diagnosing and treating cancers of the
 CC nervous system. Neural stem cells are useful for the treatment of
 CC neurological disorders, especially for repair of accidentally induced
 CC trauma in the CNS or for correction of congenital or pathological
 CC diseases of the CNS. A patient with a neurological disorder can act as a
 CC self-donor. Cells may be isolated from the patient and either sorted to
 CC extract neuroblasts, or treated in order to differentiate neuroblasts,
 CC from specific or general precursors.
 XX
 SQ Sequence 1542 BP; 238 A; 565 C; 560 G; 179 T; 0 other;
 Query Match 97.6%; Score 488; DB 20; Length 1542;
 Best Local Similarity 99.8%; Pred. NO. 7.8e-68;
 Matches 499; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCGGCGCTATGCTCCAGGCGCTCTCTCGGCGGTGATACCCGCGACCGCGCG 60
 DB 1 CCGGCGCTATGCTCCAGGCGCTCTCTCGGCGGTGATACCCGCGACCGCGCG 59
 QY 61 ATGTACAGCATGATGATGAGACCGACTCTCGCGCGCGCGCGCGCGCGCG 120
 DB 61 ATGTACAGCATGATGATGAGACCGACTCTCGCGCGCGCGCGCGCGCGCG 119
 QY 121 AACCTCGGCG 180
 DB 121 AACCTCGGCG 180

FT	variation	/note="replace(322..324, "cgn")"
FT		328..330
FT		/*tag= u
FT	variation	/note="replace(328..330, "cun")"
FT		361..363
FT		/*tag= v
FT	variation	/note="replace(361..363, "cgn")"
FT		367..369
FT		/*tag= w
FT	variation	/note="replace(367..369, "cgn")"
FT		370..372
FT		/*tag= x
FT	variation	/note="replace(370..372, "cgn")"
FT		385..387
FT		/*tag= y
FT	variation	/note="replace(385..387, "cun")"
FT		388..390
FT		/*tag= z
FT	variation	/note="replace(388..390, "cun")"
FT		406..408
FT		/*tag= aa
FT	variation	/note="replace(406..408, "agy")"
FT		409..411
FT		/*tag= ab
FT	variation	/note="replace(409..411, "cun")"
FT		421..423
FT		/*tag= ac
FT	variation	/note="replace(421..423, "cun")"
FT		424..426
FT		/*tag= ad
FT	variation	/note="replace(424..426, "cun")"
FT		505..507
FT		/*tag= ae
FT	variation	/note="replace(505..507, "cgn")"
FT		508..510
FT		/*tag= af
FT	variation	/note="replace(508..510, "cun")"
FT		513..515
FT		/*tag= ag
FT	variation	/note="replace(513..515, "agy")"
FT		583..585
FT		/*tag= ah
FT	variation	/note="replace(583..585, "agy")"
FT		631..632
FT		/*tag= ai
FT	variation	/note="replace(631..632, "cun")"
FT		676..678
FT		/*tag= aj
FT	variation	/note="replace(676..678, "cgn")"
FT		742..744
FT		/*tag= ak
FT	variation	/note="replace(742..744, "cgn")"
FT		760..762
FT		/*tag= al
FT	variation	/note="replace(760..762, "cun")"
FT		769..771
FT		/*tag= am
FT	variation	/note="replace(769..771, "agy")"
FT		778..780
FT		/*tag= an
FT	variation	/note="replace(778..780, "agy")"
FT		784..786
FT		/*tag= ao
FT	variation	/note="replace(784..786, "agy")"
FT		799..801
FT		/*tag= ap
FT	variation	/note="replace(799..801, "agy")"
FT		805..807
FT		/*tag= aq
FT	variation	/note="replace(805..807, "agy")"
FT		811..813
FT		/*tag= ar
FT	variation	/note="replace(811..813, "agy")"

[illegible]

```
DB 181 UCGUCNAGRGNCARAGARARAUGCNCARGAARAAYCCNARAUCCAYAVUCCNARG 240
QY 301 ATCAGCAAGGCGCTGGGGCGGAGTGCATGTCCTCCGAGCGCCGAGAACCGCGCTTC 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 AUHUCNAARGRURUGGNGARUCGAAKRUNUUGCNGARGNGARNAARARCCNNUY 300
QY 361 ATCGAGAGGCCAGCGGCTGCGCTGCACATGAAGAGCACCGGATTAAGATAC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 AUHGAYGARGCNAARARURBRAGRCNURCAYAUAGARARCAVCCNGAYUAYARAY 360
QY 421 CGGCGCGCGCCGCAAGCAGACGCTGCTCAAGAGGACAAGTACTGCTGCGCGCGG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 AGCCCAAGRARGRACNARACNUURUARAARAGAYARUAYUUCNURGCGNGNGN 420
QY 481 CTCCTGCGCGCGCGCGCGG 500
    : ||| ||| ||| |||
DB 421 UURURGCGNGCGNGCGNGG 440

RESULT 6
AADI1110
ID AADI1110 standard; DNA; 1085 BP.
XX
XX AADI1110;
XX
XX 24-SEP-2001 (first entry)
XX
XX Human small cell lung cancer associated gene, SOX2.
DE
XX
XX Human; small cell lung cancer; therapy; hCAP; nucleic acid; NA;
KW melanoma; cancer; colon; breast; head; neck; transitional cancer;
KW leiomyosarcoma; synovial sarcoma; cytostatic; SOX2; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 13..966
XX /tag= a
XX /product= "Human SOX2 protein"

MO20015349-A2.
PD 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US02015.
PR
XX 21-JAN-2000; 2000US-0489101.
PR
XX (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
PA (CORR ) CORNELL RES FOUND INC.
XX
XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y.
XX
XX WPI: 2001-457597/49.
DR P-PSDB; AAEO5810.
XX
XX Isolated polypeptide, used to treat or prognose a disorder
XX characterized by expression of a hCAP e.g. cancer, is encoded by an
XX isolated nucleic acid comprising an NA Group 3 or 4 molecule -
XX
XX Claim 57; Page 87; 152pp; English.
XX
XX The invention relates to nucleic acids and encoded polypeptides which
XX are cancer associated antigens expressed in patients afflicted with
XX small cell lung cancer. The molecules provided by the invention can be
XX used in the diagnosis, monitoring, research or treatment of conditions
XX characterised by the expression of one or more cancer associated
XX antigens. The polypeptide is used to treat a disorder characterised by
XX expression of a hCAP, and determine regression, progression or onset
XX of a condition characterised by expression of an abnormal amount of a
XX protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
XX are small and non-small cell lung cancer, melanoma, colon, breast, head
```

```
CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
CC The present sequence is a small cell lung cancer associated gene
CC designated as NY-SCLC-1, encoding human Sox2 protein.
XX
XX SQ Sequence 1085 BP; 253 A; 348 C; 348 G; 136 T; 0 other;

Query Match 52.9%; Score 264.4; DB 22; Length 1085;
Best Local Similarity 84.2%; Pred. No. 4, 1e-33;
Matches 298; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 143 CGGGCGGCGCGGCGGAGCGGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 CGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 124
QY 203 AGGACCGGGTCAACAGCGGCCCATGAAACGCTTCATGTGTGTGTCCCGGCGACGGCGCA 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 125 CGGACCGGGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 184
QY 263 AGATGGCCCGAGAGAACCCCAAGATGCACACTCGAGATCAGCAAGCGCGTGGGCGCG 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 185 AGATGGCCCGAGAGAACCCCAAGATGCACACTCGAGATCAGCAAGCGCGTGGGCGCG 244
QY 323 AGTGAAGGTGATGTGTCGAGGCGGAGAACGCGCGCTTCATGCGAGAGCGCGCGCTGC 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 245 AGTGAAGGTGATGTGTCGAGGCGGAGAACGCGCGCTTCATGCGAGAGCGCGCTGC 304
QY 383 GCGCGCTGCACATGAAAGAGCACCGCGGATTACAGTACCGCGCGCGCGCGAGACCAAGA 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 305 GAGCGCTGCACATGAAAGAGCACCGCGGATTAAATACCGCGCGCGCGCGCGCGCGCG 364
QY 443 CGCTGCTCAAGAGAGCAAGTACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 365 CGCTCATGAAGAGATTAAGTACACGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 418

RESULT 7
AAK52388
ID AAK52388 standard; cDNA; 1087 BP.
XX
XX AAK52388;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 933.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX MO200157190-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 05-FEB-2001; 2001WO-US04098.
PE
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-064936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
```


RESULT 9
AAK52284
ID AAK52284 standard; cDNA; 1126 BP.
AC XX
AAK52284;
DT 06-NOV-2001 (first entry)
XX XX
Human polynucleotide SEQ ID NO 829.
XX XX
Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KV tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX XX
Homo sapiens.
OS
MO200157190-AZ.
PN
XX XX
09-AUG-2001.
PD
XX XX
05-FEB-2001; 200IWO-US04098.
PF
XX XX
03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX XX
(HYSE-) HYSEQ INC.
PA
PI Tang YT, Liu C, Dirmnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI: 2001-476283/51.
XX P-PSDB; AAM79151.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
Claim 1; Page 2766-2767; 6221pp; English.
XX
The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1126 BP; 262 A; 361 C; 364 G; 139 T; 0 other;

Query Match 52.9%; Score 264.4; DB 22; Length 1126;
Best Local Similarity 84.2%; Pred. No. 4.1e-33;
Matches 298; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

OY 143 CGGCGGGGCGGAGCGGGGCGGCGGCGGCGGCGCAAGGCCAAC 202
 ||||| ||||| ||| ||| ||||| | || || ||
DB 98 CGGGGGGGCGGCGGCAACTCCACCAGCGGCGGCGGCAACCAAAAAACAGGC 157
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 203 AGGACGGGTCAACCGGCCATGAACGCTTCATGTTGTGTGTCCTCCGGGCGCAGCGCGCA 262
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db	158	CGGACCGCGTCACAGCGGCCCATGATGATCGTTCATGRTGTGGTCCCGCGGGACAGCGGGCA	217
QY	263	AGATGCGCCAGGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGGCGC	322
Db	218	AGATGGGCCCGAGGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGGCGC	277
QY	323	AGATGAGGAGCATGATCGGAGGCGGAGAAAGCGCGCTTCATGACAGAGGCCAAAGCGCTGC	382
Db	278	AGTGGCAACTTTTGTGCGAGACGGAGAGAGCGCGCTTCATGACAGAGGCTAAAGCGCTGC	337
QY	383	GCGCGTCGCACATGAAGAGCACCCGGATTACAAAGTACCGGCGCGCGCGCAAGACCAAGA	442
Db	338	GAGCGCTGCACATGAAGAGCACCCGGATTATATAATACCGGCGCGCGGGAAGAACCAAGA	397
QY	443	CGTGGTCAGAGGAGCAACTACTCGTGGCGCGCGGGCTCTTGGCGCGCGCGC	496
Db	398	CGCTCATGAAGAGATAGTATCACGCTGCCGCGGGCTGCTGGGCCCGCGCGC	451
RESULT 10			
AAK53372			
ID	AAK53372	standard; cDNA; 1501 BP.	
AC	AAK53372;		
DT	06-NOV-2001	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 2901.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; cell therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
XX	nervous system disorder; arthritis; inflammation; ss.		
OS	Homo sapiens.		
XX			
PN	WO200157190-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US04098.		
XX			
PR	03-FEB-2000; 2000US-0496914.		
PR	27-APR-2000; 2000US-0560875.		
PR	20-JUN-2000; 2000US-0598075.		
PR	19-JUL-2000; 2000US-0620325.		
PR	01-SEP-2000; 2000US-0654936.		
PR	15-SEP-2000; 2000US-0663561.		
PR	20-OCT-2000; 2000US-0693325.		
PR	30-NOV-2000; 2000US-0728422.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;		
PI	Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;		
XX			
DR	WPI: 2001-476283/51.		
DR	P-PSDB; AAM80239.		
XX			
PT	Nucleic acids encoding polypeptides with cytokine-like activities,		
PT	useful in diagnosis and gene therapy -		
XX			
PS	Claim 1; Page 5070; 6221pp; English.		
XX			
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the		
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		

CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAB0020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1501 BP; 391 A; 428 C; 450 G; 220 T; 12 other;
Query Match 52.9%; Score 264.4; DB 22; Length 1501;
Best Local Similarity 84.2%; Pred. No. 4e-33;
Matches 298; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
OY 143 CGGCGCGGGGGGGGAGGAGCGGGGGGGGGGGGGGGGGGGGGGGGCAAGGCCAAC 202
DB 91 CGGGGGGGGGGGGGGCACTCCACCGCGGGCGGGCGGCAACGAAAGAAAGGCC 150
OY 203 AGACCGGGGTCAACGGCCCATGTAACGCTTCATGTGTGTGTCGGCGGCGAGCGCGCA 262
DB 151 CGGACCGCGTCAACGGCGCCCATGTAATGCTTCATGTGTGTGTCGGCGGCGAGCGCGCA 210
OY 263 AGATGGCGCCAGGAAACCCCAAGATGCACACTCGAGATCAGCAAGCCCTGGGGCGCG 322
DB 211 AGATGGCGCCAGGAAACCCCAAGATGCACACTCGAGATCAGCAAGCCCTGGGGCGCG 270
OY 323 AGTGAAGTGTATGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 382
DB 271 AGTGAAGTGTATGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330
OY 383 GCGCGGTGCACATGAAGAGACCGCGGATTAACAGTACCGGCGCGCGCGCAAGACCAAGA 442
DB 331 GAGCGTGCACATGAAGAGACCGCGGATTAACAGTACCGGCGCGCGCGCAAGACCAAGA 390
OY 443 GCGGTGCACAGAAAGACAGTACTGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 496
DB 391 GCGTCAAGAAAGATTAAGTACAGCGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 444
RESULT 11
ABI99318
ID ABI99318 standard; cDNA; 2283 BP.
XX
AC ABI99318;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:179.
XX
KW Mouse; Ischemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN M0200188188-A2.
XX
PD 22-NOV-2001.
XX
PE 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
XX (UNINT-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PA Ishikawa K, Asai S, Takehashi Y, Nagata T, Ishii Y;
PI WPI; 2002-034733/04.
DR P-PSDB; ABB57082.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
XX genes -
XX

PS Claim 2; Page 518-521; 2690pp; English.
XX
XX The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischemia, occlusive ischemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912), encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 2283 BP; 597 A; 583 C; 614 G; 488 T; 1 other;
Query Match 52.6%; Score 262.8; DB 24; Length 2283;
Best Local Similarity 82.9%; Pred. No. 6.7e-33;
Matches 300; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
OY 137 CGGGGGGGGGCGCGCGGGCGGAGCGGGGGCGGCGGCGGGGGCGGCGCAAG 196
DB 414 CGGGGGGGGGGGGGGAGGAGGAGCAACCGCGCGCGCGCGGCAACGAGAA 473
OY 197 CCAACGAGACCGGGTCAACCGCCCATGAAGCGCTTCATGTGTGTGTCGCGGGCAGC 256
DB 474 ACAGCCCGGACCGCGTCAAGAGGCCCATGAAGCGCTTCATGTGTGTGTCGCGGGCAGC 533
OY 257 GGGCAGATGGGCGCCAGAGAACCCCAAGATGCACACTCGGAGATCAGCAAGCGCCGCGG 316
DB 534 GCGCTTAAGATGGCCAGAGAACCCCAAGATGCACACTCGGAGATCAGCAAGCGCCGCGG 593
OY 317 GGGCGGAGTGAAGTATGTCGAGGCGGAGAAAGCGGCGCTTCATCAGCAGAGCCAAAGC 376
DB 594 GCGCGAGTGAAGACTTTTGTCCGAGACCGAGAAAGCGGCGCTTCATCAGCAGAGCCAAAGC 653
OY 377 GCGTGGCGGGCGTGCACATGAAGAGACCGCGGATTAACAGTACCGGCGCGCGCGCAAGA 436
DB 654 GCGTGGCGGGCGTGCACATGAAGAGACCGCGGATTAACAGTACCGGCGCGCGCGCAAGA 713
OY 437 CCAAGACGTGCTCAAGAGAGCAAGTACTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGG 496
DB 714 CCAAGACGTGCTCAAGAGAGATTAAGTACAGCTTCCGGAGGCTTGCTGCGGCCCGCGGCG 773
OY 497 CG 498
DB 774 GG 775
RESULT 12
AAD1118
ID AAD1118 standard; DNA; 2509 BP.
XX
AC AAD1118;
XX
DT 24-SEP-2001 (first entry)
XX
DE Human small cell lung cancer associated gene, SOX3.
XX
KW Human; small cell lung cancer; therapy: hCAP; nucleic acid; NA;
KW melanoma; cancer; colon; breast; head; neck; transitional cancer;
KW leiomyosarcoma; synovial sarcoma; cytosolic; SOX3; ds.
XX
XX Homo sapiens.
XX
XX Key location/Qualifiers
FH CDS 441..1773
FT /*tag= a
FT /product= "Human SOX3 protein"
XX


```
QY 213 CAACGGCC-ATGACGCTTCATGNGTGTCGCCGGGAGCGGCAAGATGGCC 271
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 GAAAGCGCCCAATGAAACGCTTCATGTGTGTGTCGCCGGGCAAGGCAAAATGGCC 900
QY 272 AGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGGCCGAGTGAAG 331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 CGGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGGCCGAGTGAAG 960
QY 332 TCATGTCGGGCGGCAAGCGGCGCTTCATCGACGAGGCGCAAGCGGCTGGCGCTGC 391
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 TGCTGACCGGCGGCAAGCGGCGCTTCATCGACGAGGCGCAAGCGCTTCGCGCTGC 1020
QY 392 ACATGAAGAGACCGCGGATTCACAACTACGCGGCGGCGGCAAGCAAGCGCTGCA 451
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 ACATGAAGAGATTCGCGACTACAACTACGCGGCGGCGGCAAGCAAGCGCTGCTCA 1080
QY 452 AGAAGCAAGTACTCGCTGGCGGCGGCTCTGTCGGCGGCGGCGG 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 AGAAAGATAGTACTCGCTGCCCGCGGCTCTGCTCGCGTCCG 1128
```

RESULT 14

```
AD11119
ID AD11119 standard; DNA; 8372 BP.
```

```
XX AAD11119;
```

```
DT 24-SEP-2001 (first entry)
```

```
DE Human small cell lung cancer associated gene, SOX21.
```

```
XX Human; small cell lung cancer; therapy; hCAAP; nucleic acid; NA;
```

```
KW melanoma; cancer; colon; breast; head; neck; transitional cancer;
```

```
KW leiomyosarcoma; synovial sarcoma; cytostatic; SOX21; ds.
```

```
XX Homo sapiens.
```

```
OS Key Location/Qualifiers
```

```
FT CDS 1167..1997
```

```
FT FT /tag= a
```

```
FT product= "Human SOX21 protein"
```

```
XX MO20015349-A2.
```

```
XX 26-JUL-2001.
```

```
XX 19-JAN-2001; 2001WO-US02015.
```

```
XX 21-JAN-2000; 2000US-0489101.
```

```
XX (LUDW-) LUDWIG INST CANCER RES.
```

```
XX PA (SLOK ) SLOAN KETTERING INST CANCER RES.
```

```
XX PA (CORR ) CORNELL RES FOUND INC.
```

```
XX PI Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
```

```
XX DR MPI: 2001-457597/49.
```

```
XX P-PSDB; AAE05814.
```

```
XX Isolated polypeptide, used to treat or prognose a disorder
```

```
XX characterized by expression of a hCAAP e.g. cancer, is encoded by an
```

```
XX isolated nucleic acid comprising an NA Group 3 or 4 molecule -
```

```
XX Claim 57; Page 100-105; 152pp; English.
```

```
CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
CC are small and non-small cell lung cancer, melanoma, colon, breast, head
CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
CC The present sequence is a small cell lung cancer associated gene
CC designated as NY-SCIC-10, encoding human SOX21 protein.
```

```
XX Sequence 8372 BP; 1996 A; 2143 C; 2104 G; 2124 T; 5 other;
```

Query Match

```
Best local similarity 45.7%; Score 228.6; DB 22; Length 8372;
Matches 318; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
```

```
QY 33 GGTGCGGTGAACCCCGCACCGCCCGGATGATACGATATGATGAGACCGACCTGCA 92
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1010 GGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1069
QY 93 CTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 152
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1070 GCGGAGCGGCTCTCCGGGCGAGCAACATTTGATTTCTCCGGGCGGCGGCGGCGGCGG 1129
QY 153 CGGGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 212
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1130 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1189
QY 213 CAAGGCGCGGATGAGCGCTTCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 272
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1190 CAAGGCGCGGATGAGCGCTTCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1249
QY 273 GGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGGCGGCGGCGGCGG 332
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1250 GGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGGCGGCGGCGGCGG 1309
QY 333 CATGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 392
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1310 GCTCAGCAGTCTGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1369
QY 393 CATGAGGAGCACCGGATTCACAACTCGAGATCAGCAAGCGCTGGGGCGGCGGCGGCGG 452
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1370 CATGAGGAGCACCGGATTCACAACTCGAGATCAGCAAGCGCTGGGGCGGCGGCGGCGG 1429
QY 453 GAGGACAACTACTCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 499
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1430 GAGGACAACTACTCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1476
```

RESULT 15

```
AAS61625
ID AAS61625 standard; cDNA; 412 BP.
```

```
XX AAS61625;
```

```
XX 29-JAN-2002 (first entry)
```

```
XX Lung small cell carcinoma antigen, cDNA #166.
```

```
KW Human; cytostatic; antitumour; lung small cell cancer antigen;
```

```
KW tumour; lung cancer; ss.
```

```
XX Homo sapiens.
```

```
XX WO200177168-A2.
```

```
XX 18-OCT-2001.
```

```
XX 11-APR-2001; 2001WO-US11859.
```

```
XX 11-APR-2000; 2000US-196780P.
```

```
XX 21-JUN-2000; 2000US-213361P.
```

```
XX 01-SEP-2000; 2000US-229763P.
```

```
XX 05-SEP-2000; 2000US-230629P.
```

```
XX 14-SEP-2000; 2000US-232659P.
```

```
XX 19-DEC-2000; 2000US-257037P.
```

```
XX 08-JAN-2001; 2001US-260796P.
```

XX
PA (CORI-) CORIXA CORP.
XX

PI Lodes MJ, Wang T, Mohamach R, Indirias CY;
XX

DR WPI; 2002-010896/01.
XX

XX Lung tumour polynucleotide and polypeptides useful in therapy and
PT diagnosis of cancer especially lung cancer -
XX

PS Claim 1; Page 180; 295pp; English.
XX

XX The invention relates to novel isolated lung small cell cancer antigen
CC polynucleotides (I) and polypeptides (II) used in a method of detecting
CC cancer in a patient. The method is optionally performed by
CC utilising oligonucleotides (III), where the biological sample
CC from the patient is contacted with (III), detecting the amount of
CC polynucleotide hybridised to (III) in the sample and comparing the
CC amount of polynucleotide to a predetermined cut-off value and thereby
CC determining cancer in a patient. (I), (II) or antigen-presenting cells
CC expressing (II) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein. The method comprises contacting T cells
CC with one of the components under conditions to permit the stimulation
CC and/or expansion of the cells. A composition comprising (I) is useful for
CC stimulating an immune response in a patient and for inhibiting the
CC development of a cancer especially lung cancer in a patient. An
CC isolated T cell population is useful for removing tumour cells from the
CC biological sample and for inhibiting the development of cancer in a
CC patient. AAS61460-AAS61874 represent novel human lung small cell
CC cancer antigen coding sequences of the invention.
XX

SQ Sequence 412 BP; 83 A; 133 C; 144 G; 51 T; 1 other;

Query Match 44.2%; Score 220.8; DB 24; Length 412;

Best Local Similarity 84.1%; Pred. No. 2.7e-26; Indels 0; Gaps 0;

Matches 249; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 204 GGACGGGTCAACGCGCCATGAAACGCTTCATGTGTGTCCCGCGGCAAGCGCGCAA 263

DB 15 GGACGACGTCAAGCGGCCCATGACGCTTCATGTGTGTCCCGCGGCTGACGCGCGCAA 74

QY 264 GATGCCCCAGGAAACCCCAAGATGCAACAATCGAGATCAGCAAGCGCTGGGGGCCGA 323

DB 75 GATGCCCCAGGAAACCCCAAGATGCAACAATCGAGATCAGCAAGCGCTGGGGGCCGA 134

QY 324 GTGGAAGGTCATGTCGAGCGGAGAGCGCGCTTCATGACGAGGCGCAAGCGGCTGCG 383

DB 135 GTGGAAGGTCATGTCGAGCGGAGAGCGCGCTTCATGACGAGGCGCAAGCGGCTGCG 194

QY 384 CGCGCTGCATGAAGGAGACCCCGGATTACAAGTACCGGCGCGCCCAAGACCAAGAC 443

DB 195 CGCATGACATGAAGGAGACCCCGGATTACAAGTACCGGCGCGCCCAAGACCAAGAC 254

QY 444 GCTGCTCAAGAAAGACAACTACTGCTGGCGGCGGCGCTCTGGCGGCGGCGCGG 499

DB 255 GCTGCTCAAGAAAGACAACTACTGCTGGCGGCGGCGCTCTGGCGGCGGCGGCGG 310

Search completed: October 10, 2002, 14:28:39

Job time : 91 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 18.7333 Seconds

(Without alignments)
6556.054 Million cell updates/sec

Title: US-09-489-101a-3_COPY_1_500

Perfect score: 500
1 cacacagcccgcatgtacaa.....atgacagttacgcgcacat 500Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	140	28.0	3923	3	US-08-860-635A-20
2	140	28.0	3923	4	US-09-281-476-20
3	134.6	26.9	1661	1	US-08-196-016A-1
4	129.8	26.0	1667	1	US-08-196-016A-2
5	129.4	25.9	2249	3	US-08-860-635A-18
6	129.4	25.9	2249	4	US-09-281-476-18
7	99.8	20.0	3243	2	US-08-239-276-6
8	99.8	20.0	3243	2	US-08-468-579B-6
9	99.8	20.0	3243	2	US-08-468-579B-6
10	89.4	17.9	5051	2	US-08-239-276-2
11	89.4	17.9	5051	2	US-08-468-579B-2
12	89.4	17.9	5051	2	US-08-468-579B-2
13	80.8	16.2	1570	2	US-08-239-276-5
14	80.8	16.2	1570	2	US-08-468-579B-5
15	80.8	16.2	1570	2	US-08-468-579B-5
16	73.6	14.7	1397	2	US-08-239-276-1
17	73.6	14.7	1397	2	US-08-468-579B-1
18	73.6	14.7	1397	3	US-08-468-579B-1
19	62.6	12.5	684	4	US-08-998-416-672
20	49	9.8	7218	1	US-08-332-463-14
21	44.6	8.9	2040	2	US-08-821-355A-1
22	44.6	8.9	2040	2	US-09-003-687A-1
23	44.6	8.9	2040	3	US-09-136-605-1
24	44.6	8.9	2444	2	US-08-821-355A-2
25	44.6	8.9	2444	2	US-09-003-687A-2
26	44.6	8.9	2444	2	US-09-136-605-2
27	44.6	8.9	23673	4	US-09-773-816-1

28	43	8.6	1898	1	US-07-814-964-8	Sequence 8, Appl
29	43	8.6	1898	1	US-08-258-442-8	Sequence 8, Appl
30	43	8.6	1898	1	US-08-328-809-3	Sequence 3, Appl
31	43	8.6	1898	5	PCT-US92-11107-8	Sequence 6, Appl
32	43	8.6	2839	1	US-07-814-964-6	Sequence 6, Appl
33	43	8.6	2839	1	US-08-258-442-6	Sequence 6, Appl
34	43	8.6	2839	1	US-08-328-809-1	Sequence 1, Appl
35	43	8.6	2839	3	US-09-015-003-1	Sequence 1, Appl
36	43	8.6	2839	5	PCT-US92-11107-6	Sequence 6, Appl
37	42.8	8.6	2032	4	US-09-045-284A-1	Sequence 14, Appl
38	42.8	8.6	4722	4	US-08-979-608A-14	Sequence 9, Appl
39	41.6	8.3	1444	1	US-07-814-964-9	Sequence 9, Appl
40	41.6	8.3	1444	1	US-08-258-442-9	Sequence 4, Appl
41	41.6	8.3	1444	1	US-08-328-809-4	Sequence 9, Appl
42	41.6	8.3	1444	5	PCT-US92-11107-9	Sequence 5, Appl
43	41.4	8.3	2277	1	US-08-676-967-5	Sequence 5, Appl
44	41.4	8.3	2277	1	US-08-676-974-5	Sequence 5, Appl
45	41.4	8.3	2277	2	US-09-098-487-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-860-635A-20
Sequence 20, Application US/08860635A
Patent No. 6143878
GENERAL INFORMATION:
APPLICANT: Koopman, Peter
APPLICANT: Goodfellow, Peter
TITLE OF INVENTION: SOX-9 GENE AND PROTEIN AND
TITLE OF INVENTION: USE IN THE REGENERATION OF BONE OR CARTILAGE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,635A
FILING DATE: 29-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM9714
FILING DATE: 29-NOV-1994
APPLICATION NUMBER: AU PM9835
FILING DATE: 05-DEC-1994
APPLICATION NUMBER: PCT/AU95/00799
FILING DATE: 29-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10981
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3923 base pairs
TYPE: nucleic acid
TOPOGENESIS: single
STRADEDNESS: linear
MOLECULE TYPE: CDNA
US-08-860-635A-20


```

: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEITICAL: NO
: ANTI-SENSE: NO
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: Y-chromosome, SRY gene
US-08-196-016A-1

```

Query Match	26.98;	Score 134.6;	DB 1;	Length 1661;
Best Local Similarity	69.78;	Pred. No. 2.1e-21;		
Matches 182;	Conservative 0;	Mismatches 79;	Indels 0;	Gaps 0;

Accession	Sequence	Position
Oy	114 GAAAAACAGCCCGGACCGGTGAAGGGCCCAATGATGCTTATGATGGTGTCCCGG	173
Db	1052 GGAGAGCAGCCAGGACCAAGTAAAGCAACCCATGAACGCTTATGTGTGTCTGTGA	1111
Oy	174 GCAGCGGCGCAAGATGGCCGACAGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGC	233
Db	1112 ACGAAGACGAAAGGTGGCTCTAGAGAAATCCCAAAATGAAAAACCTCAGACATCAGCAAGCA	1171
Oy	234 CCTGGGCGCCGAGTGGAAATTTTGGCGAGACGGAGAGCGGCCGCTTCATCGCAGAGC	293
Db	1172 GCTGGGATATGTGTGAAAAAGGCTTACAGATGCTGAAAAGCGCCATTTCTTTGAGGAGGC	1231
Oy	294 TAAAGCGGTGCGAGCGCTGCACATGAAGAGACCCCGATTAATTAATACGGCCCCGGCG	353
Db	1232 ACAGAGACTACTAGCCATACACCGAGACAAATACCGGGCTTAATAATATGACCTCTGCG	1291
Oy	354 GAAACCAAGAGCTCATGAA 374	
Db	1292 GAGAGCCAGAGGCCACAGAA 1312	

RESULT 4

US-08-196-016A-2

; Sequence 2, Application US/08196016A
; Patent No. 5596089

```

1  GENERAL INFORMATION:
2
3  APPLICANT: SILVERSIDES, David W.
4
5  APPLICANT: DANEAU, Isabelle M.F.
6
7  APPLICANT: HOUDE, Alain
8
9  TITLE OF INVENTION: SRY GENOMIC SEQUENCES AND USE IN TISSUE
10
11 TITLE OF INVENTION: SEXING
12
13 NUMBER OF SEQUENCES: 34
14
15 CORRESPONDENCE ADDRESS:
16
17 ADDRESSEE: KLABER & JACKSON
18
19 STREET: Continental Plaza, 411 Hackensack Avenue
20
21 CITY: Hackensack
22
23 STATE: N.J.
24
25 COUNTRY: U.S.A.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: 05/0
 ; FILING DATE: 14-FEB-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:

NAME: JACKSON, David A.
REGISTRATION NUMBER: 26,742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521

```

?      TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?      HYPOTHEetical: NO
?      ANTI-SENSE: NO
?      POSITION IN GENOME:
?      CHROMOSOME/SEGMENT: Y-chromosome, SRY gene
?      US-08-136-016A-2

```

Query Match	26.0%;	Score 129.8;	DB 1;	Length 1667;
Best Local Similarity	70.6%;	Pred. No. 2.4e-20;		
Matches 173; Conservative	0;	Mismatches 72;	Indels 0;	Gaps 0.

QY	122	GGCGAGCCGCTCAAGCGCGCCATGAATGCTTCATGGTGTGATCCCGCGGAGCGGC	181
Db	791	GCCAGATCGTGTCAAGCCACCCATGACGCTTTCATGTGTGTGTCTCGTGATCAAAAGA	850
QY	182	GCAGATGCGCCAGGAGAAACCCAGATGCACAATCGGATGACGACGCGCTGGGCG	241
Db	851	GAAAGTGCGCTAGAGAAACCTCCAATGCAAACTGAGATCAGCAAGTGGTGGAAT	910
QY	242	CCGAGTGGAAACTTTGTGGGAGACGAGAAAGCGGCGTTCATGACGAGCGCTTAAGCGC	301
Db	911	GCAAGTGGAAATGCTTTACAGAAAGCCGGAAGAGCGCCCATCTTCTTGAGAGGACCAAGAGCG	970
QY	302	TTCGAGCGCTGCACATGAAGAGACACCGGATTTATAATACCGGCCCGCGCGAAAAACA	361
Db	971	TTCAGCGGCTGCACGAGATTAATACCGGCGCTATAATACCGACCTCTTCGCAAGGAG	1030
QY	362	AGACG 366	
Db	1031	AGAGG 1035	

RESULT 5

US-08-860-635A-18

; Sequence 18, Application US/08860635A
; Patent No. 6143878

GENERAL INFORMATION:
APPLICANT: Koopman, Peter
APPLICANT: Goodfellow, Peter
TITLE OF INVENTION: SOX-9 GENE AND PROTEIN AND
TITLE OF INVENTION: USE IN THE REGENERATION OF BONE OR CARTILAGE
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: U.S.A.
ZIP: 11530

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
;

```

```

;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
;

```

```

; SOFTWARE: FastSEQ Version 1.5
;
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/860,635A
FILING DATE: 29-MAY-1997

CLASSIFICATION: 514
PRIOR APPLICATION DATA

PRIOR APPLICATION DATA.
 APPLICATION NUMBER: AU PM9714
 FILING DATE: 29-NOV-1994

```

; FILING DATE: 29-NOV-1994
; APPLICATION NUMBER: AU PM9835
; FILING DATE: 05-DEC-1994

```

FILING DATE: 05-DEC-1994
APPLICATION NUMBER: PCT/AU95/007995

; FILING DATE: 29-NOV-1995
 ;
 ; ATTORNEY/AGENT INFORMATION:
 ;

FILING DATE: 17-FEB-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: KURT G. BRISCOE
 REGISTRATION NUMBER: 33,141
 REFERENCE/DOCKET NUMBER: MDI 251.4-KGB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 332-1700
 TELEFAX: (914) 332-1844
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3243 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match Similarity	20.0%	Score 99.48	DB 2	Length 3243
Best Local Similarity	62.8%	Pred. No. 1e-13		
Matches 155; Conservative	0	Mismatches 92	Indels 0	Gaps 0

Qy	129	CCGGTCAACGGCCCATGAATGCTTTCATGCTGTGTGTCCTCCGGGGCAGCGCGCAGAT	188
Db	1015	CCACATCAAGAGGCCCATGAACCCCTTTCATGTGTGTGGCCCAAGGATGAGCGGAGGAAGAT	1074
Qy	189	GGCCCAAGAGAACCCCAAGATGCACAATCTCGAGATCAGCAAGCGGCTGTGGGCGCGAGTG	248
Db	1075	CTGTCAGACGCTTCTCCAGACACTGCACAACCTCAGCATCAACCAAGATCTCTTGATCTCGCTG	1134
Qy	249	GAACCTTTTCTCGAGAGCAGGAAGCGGCGCTTCATCGACGAGGCTAAAGCGGCTCGAGC	308
Db	1135	GAAGTCCATATGCCAACCAAGSAGAAAGCAAGCCCTCTCTATGAGSAAACAGCGCGCTGAGCGC	1194
Qy	309	GCTGCACTAATAGAGACACCCGGATTAATAATTCGGGCCCGCGCGGAGAAAACCAAGACGT	368
Db	1195	GCACACCACTGAGAGAAATATCTGACTACAAATCAAGACCGCGGCCCAAGCGCACCTGAT	1254
Qy	369	CATGAG 375	
Db	1255	CTGTGAG 1261	

RESULT 8
 US-08-468-579B-6
 Sequence 6, Application US/08468579B
 Patent No. 5981700
 GENERAL INFORMATION:
 APPLICANT: Rablin, Daniel
 TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGEN-
 TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sprung Kramer Schaefer & Briscoe
 STREET: 660 White Plains Road
 CITY: Tarrytown
 STATE: New York
 COUNTRY: USA
 ZIP: 10591-5144
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: System 7.5
 SOFTWARE: WordPerfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/468,579B
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/239,276
 FILING DATE: 05-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/872,646
 FILING DATE: 08-JUN-1992
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/715,181
 FILING DATE: 14-JUN-1991
 PRIOR APPLICATION DATA: US 07/441,703
 FILING DATE: 04-DEC-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/312,543
 FILING DATE: 17-FEB-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: KURT G. BRISCOE
 REGISTRATION NUMBER: 33,141
 REFERENCE/DOCKET NUMBER: MDI 251.5-KGB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 332-1700
 TELEFAX: (914) 332-1844
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3243 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match	20.0%;	Score 99.8;	DB 2;	Length 3243;
Best Local Similarity	62.8%;	Pred. No. 1e-13;		
Matches 155;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;

[illegible]

RESULT 9
 ; Sequence 6, Application US/08468577B
 ; Patent No. 6001804
 ; GENERAL INFORMATION:
 APPLICANT: Rablin, Daniel
 TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENMS
 TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Spring Krimer Schaefer & Briscoe
 STREET: 660 White Plains Road
 CITY: Tarrytown
 STATE: New York
 COUNTRY: USA
 ZIP: 10591-5144
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: System 7.5
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/468,577B
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 514


```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: KURT G. BRISCOE
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.8-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3243 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-468-577B-6

```

```

Query Match      20.0%; Score 99.8; DB 3; Length 3243;
Best Local Similarity 62.8%; Pred. No. 1e-13;
Matches 155; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```

```

QY 129 CCGCCGACGAGCCGACATGATGCTTCATGTCGTCCTCCGCGGCGACGCGCCGACAT 188
    || || || || || || || || || || || || || || || || || || || || ||
DB 1015 CCACATCAAGAGCCGACATGACGCTTCATGTCGTCGTCGCGCAAGATGACGAGAAAT 1074
    || || || || || || || || || || || || || || || || || || || || ||
QY 189 GCGCCGAGAAACCCGACATGACACACTGAGATGACGACGAGCGCTGGCGCGGAGTGTG 248
    || || || || || || || || || || || || || || || || || || || || ||
DB 1075 CCGTGAAGCCTTCCGACATGACCAACTCCAGATCAGCAAGATCTCTGATCTCGCTG 1134
    || || || || || || || || || || || || || || || || || || || || ||
QY 249 GAACTTTTGTGCGAGACGAGAGCGCGCTTCATGACGAGGCTTAAGCGCTGCGAGC 308
    || || || || || || || || || || || || || || || || || || || || ||
DB 1135 GAAGTCATGACCAACGAGAGAGAGAGCCCTACTATGAGAAACGAGCGGCGTGAAGCG 1194
    || || || || || || || || || || || || || || || || || || || || ||
QY 309 GCTGCATGAAGAGACACCGGATTTAATATACCGCGCGCGGCGAAACCAAGAGCGCT 368
    || || || || || || || || || || || || || || || || || || || || ||
DB 1195 GCAGACCGTGAAGATATCTGACTACAAAGTACAAAGCCGCGCGCAAGCGCACCTGCAT 1254
    || || || || || || || || || || || || || || || || || || || || ||
QY 369 CATGAAG 375
    || || ||
DB 1255 CGTGAAG 1261
    || || ||

```

```

RESULT 10
US-08-239-276-2
Sequence 2, Application US/08239276
Patent No. 5840836
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

```

```

COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/239,276
FILING DATE: 05-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: KURT G. BRISCOE
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.4-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5051 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-239-276-2

```

```

Query Match      17.9%; Score 89.4; DB 2; Length 5051;
Best Local Similarity 60.2%; Pred. No. 2e-11;
Matches 147; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```

```

QY 118 AACAGCCCGGAGCCGCTCAAGCGGCCCATGATGCTTCATGTCGTCCTCCGCGGCGAG 177
    || || || || || || || || || || || || || || || || || || || || ||
DB 1800 AGCAGCCGAGCCGACATTAAGGAGCCCAATGATGCTTCATGTCGTCGTCGCGCAAGT 1859
    || || || || || || || || || || || || || || || || || || || || ||
QY 178 GCGCCGAAAGATGCGCCGAGAGAACCCGACATGACACTGAGATGACGAGATCGAGCGGCTG 237
    || || || || || || || || || || || || || || || || || || || || ||
DB 1860 AGGAAAAAATCTCTGAGCGCTTCCCGGACATGATTAATCTCAACATTAAGCAAAATCTTA 1919
    || || || || || || || || || || || || || || || || || || || || ||
QY 238 GCGCCGAGTGAATCTTTGTGCGAGAGCGAGAACCGCGCTTCATGACGAGGCTAAG 297
    || || || || || || || || || || || || || || || || || || || || ||
DB 1920 GGATCTCGCTGGAATTCATATGTCCAACGAGAGAACCACTTATATGAAGAGCGAGGCC 1979
    || || || || || || || || || || || || || || || || || || || || ||
QY 298 GCGCTGCGAGCGCTGCACATGAAGAGACCCGCGATTTAATATACCGCGCGCGGAGAA 357
    || || || || || || || || || || || || || || || || || || || || ||
DB 1980 CGGCTAAGCAAGATCCACTTGAAGAGATACCAAACTATATAAACCCGACCGAGAA 2039
    || || || || || || || || || || || || || || || || || || || || ||
QY 358 ACCA 361
    || ||
DB 2040 CNCA 2043
    || ||

```

```

RESULT 11
US-08-468-579B-2
Sequence 2, Application US/08468579B
Patent No. 5981700
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York

```

```

Query Match: 17.9%; Score 89.4; DB 2; Length 5051;
Best Local Similarity 60.2%; Pred. No. 2e-11;
Matches 147; Conservative 0; Mismatches 97; Indels 0; Gaps 0.

OY 118 AACAGCCCGACCGCGCTCAAGCGGCCATGSAATGCGCTTCATGTGTGTCGCGCGGAC 177
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1800 AGCAGCGACGCCACACATTTAAAGCGACCATTGAATGATTCATGTGTTTGGCCAAAGATGAG 1859

OY 178 CGCGCGCAAGATGGCCCGAGGAGAACCCCAAGATGCACAACTGGAGATTCAGCAACCGCTG 237
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1860 AGGAGAAAATCTCTTCAGCGCTTCCCGCAGCATGCTAATCTCACAATAGAAATCTTA 1919

OY 238 GCGCGCGCAGTGGAAACTTTTGTGCGAGACGAGAAAGCGCGCTTCATTCGACGAGGCTAAG 297
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1920 GGATCTCTGCTGGAAATCAATGTCACCAACGAGGAAGCAACCTTATTAAGAGCAGGCC 1979

OY 298 CGCGTGGAGCGGTGTGACATGAAGAGACCCCGATTAATATACCGGCCCGCGGGGAA 357
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1980 CGGCTTAAGCAAGATCCACTTAGAGAAGTACCCAAATCTAATATCAAAACCCCGACGAAA 2039

OY 358 ACCA 361
      | |
Db 2040 CNCA 2043

RESULT 12
US-08-468-577B-2
; Sequence 2, Application US/08468577B
; Patent No. 6001804
; GENERAL INFORMATION:

```

[illegible]

RESULT 13

US-08-239-276-5
Sequence 5, Application US/08239276
Patent No. 5840836
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/239,276
FILING DATE: 05-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.4-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1570 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-239-276-5

Query Match 16.2%; Score 80.8; DB 2; Length 1570;
Best Local Similarity 62.9%; Pred. No. 1.3e-09;

Matches 141; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 129 CCGCGTCAGAGCGCCCAAGATGCTTCATGCTGTGTCGCCCGGAGCAAGCGGCAAGAT 188
DB 163 CCACATCAAGAGCGCCATGACGCTTCATGCTGTGCGCAAGATGACGAGGAAGAT 222
QY 189 GGGCCAGAGAACCCCAAGATGCTTCATGCTGTGTCGCCCGGAGCAAGCGGCAAGAT 248
DB 223 CCGGAGAGCGCTTCAGATGACGCTTCATGCTGTGTCGCCCGGAGCAAGCGGCAAGAT 282
QY 249 GAAACTTTTGTGCGAGAGAGAGAGAGCGGCGCTTCATGCTGTGTCGCCCGGAGCAAG 308
DB 283 GAAGTCATGACCAACAGAGAGAGAGAGCGGCGCTTCATGCTGTGTCGCCCGGAGCAAG 341
QY 309 GCTGCATGAG 352
DB 342 TCATCAGCTGAGAGAGATTCCTGACTACAGATACAGAGAGAGAGAGAGAGAGAGAGAG 385

RESULT 14

US-08-468-579B-5
Sequence 5, Application US/08468579B
Patent No. 5981700
GENERAL INFORMATION:
APPLICANT: Rabin, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,579B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.5-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1570 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-468-579B-5

Query Match 16.2%; Score 80.8; DB 2; Length 1570;
Best Local Similarity 62.9%; Pred. No. 1.3e-09;

Matches 141; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 129 CCGCGTCAGAGCGCCCAAGATGCTTCATGCTGTGTCGCCCGGAGCAAGCGGCAAGAT 188
DB 163 CCACATCAAGAGCGCCATGACGCTTCATGCTGTGCGCAAGATGACGAGGAAGAT 222
QY 189 GGGCCAGAGAACCCCAAGATGCTTCATGCTGTGTCGCCCGGAGCAAGCGGCAAGAT 248
DB 223 CCGGAGAGCGCTTCAGATGACGCTTCATGCTGTGTCGCCCGGAGCAAGCGGCAAGAT 282
QY 249 GAAACTTTTGTGCGAGAGAGAGAGAGCGGCGCTTCATGCTGTGTCGCCCGGAGCAAG 308
DB 283 GAAGTCATGACCAACAGAGAGAGAGAGCGGCGCTTCATGCTGTGTCGCCCGGAGCAAG 341
QY 309 GCTGCATGAG 352

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 695.733 Seconds

(without alignments)
9699.805 Million cell updates/sec

Title: US-09-489-101a-3_COPY_1_500

Perfect score: 500
Sequence: 1 caccagccgccgcatgtacaa.....atggacagttacgcacat 500

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estov:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_hiv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	580	10	BF939816 nacc81b12.
2	500	100.0	694	10	BF917204 B1917204 603181352
3	499	99.8	509	9	A1094794 ga17f03.x
4	499	99.8	537	9	AM071909 w56c10.x
5	499	99.8	552	9	AM131791 x134f05.x
6	498	99.6	537	9	A1971611 w706a10.x
7	496.8	99.4	641	9	BE206866 ba04g09.y
8	496.8	99.4	728	10	BE672418 7a0e05.x
9	496.4	99.3	507	9	A1571299 tm44e04.x
10	495.4	99.1	537	9	A1480221 tm7g05.x
11	488	97.6	965	9	A1356682 qy22a10.x
12	484.4	96.9	607	9	AM055151 w201d08.x
13	479	95.8	485	9	AM028031 w563a01.x
14	477	95.4	514	9	A1566261 tq70d10.x
15	467	93.4	740	10	BE792833 601584834
16	465	93.0	764	10	BE827056 602749106
17	465	93.0	898	10	BE305566 601889087

18	463	92.6	834	10	BE797259	BE797259 601587618
19	463	92.6	861	10	BE795592	BE795592 601589105
20	462.4	92.5	465	9	A1096977	qba2c12.x
21	459.4	91.9	465	9	A1146367	qba1911.x
22	459	91.8	460	9	AM005368	w285h12.x
23	454.2	90.8	460	9	A1292258	qm76c01.x
24	453	90.6	478	10	BF513177	BF513177 UI-R-BW1
25	452	90.4	480	9	A1146406	qa06b04.x
26	451	90.2	490	9	A1418146	tf68e09.x
27	445	89.0	483	9	A1421743	tf39b01.x
28	444	88.8	600	9	A1102567	EST211856
29	443.6	88.7	452	9	A1631443	w889a09.x
30	441	88.2	450	9	A1373018	q202g05.x
31	441	88.2	450	9	A1052267	oy58e06.x
32	441	88.2	739	10	BF305922	BF305922 601889206
33	440.2	88.0	473	9	A1990431	A1990431 w74c08.x
34	439.8	88.0	567	9	A1145897	UI-R-BT0
35	435.4	87.1	462	10	BF512088	BF512088 UI-R-BW1
36	435.4	87.1	463	9	A1199308	q14a10.x
37	435	87.0	448	9	AM131705	xf33d09.x
38	430.8	86.2	472	9	A1097136	qb60h04.x
39	430.6	86.1	517	9	A1112078	UI-R-YO-m
40	421	84.2	465	9	AM297711	UI-R-BW0
41	417.4	83.5	427	9	AM615144	bg73e05.x
42	417.4	83.5	506	9	AM533022	UI-R-BW0
43	415.4	83.1	784	10	BF305813	BF305813 60189179
44	408.2	81.6	518	9	AM044044	w205g01.x
45	408	81.6	606	10	BM083874	BM083874 ima9eqc_6

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
BF939816 nacc81b12.x1 NCI_CGAP_Brn23 Homo sapiens CDNA clone IMAGE:3440614 3'
similar to SW:SOX2_SHEEP P54231 TRANSCRIPTION FACTOR SOX-2. [1]
; contains TAR1.tl TAR1 repetitive element ;, mRNA sequence.

ACCESSION
BF939816
VERSION
BF939816.1 GI:12357136
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 580)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index

JOURNAL
COMMENT
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs.r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
infoimage.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 505.
Location/Qualifiers

FEATURES
source
1..580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3440614"
/clone_lib="NCI_CGAP_Brn23"
/tissue_type="gliodlastoma (pooled)"
/lab_host="DH10B"

R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaudo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
 Insert length: 911 Std Error: 0.00
 Seq primer: -400P from Gibco
 High quality sequence stop: 462.
 Location/Qualifiers

FEATURES

source
 1. 537
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2480730"
 /clone_lib="NCI-CGAP-GC6"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaudo."
 BASE COUNT 124 a 161 c 190 g 60 t 2 others
 ORIGIN

Query Match 99.6%; Score 498; DB 9; Length 537;
 Best Local Similarity 99.6%; Pred. No. 1.7e-80;
 Matches 498; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 CACAGCGCCCGCATGTATACATGATGAGAGCGAGCTGAAGCCCGCGCCGCGACGAA 60
 2 CACAGCGCCCGCATGTATACATGATGAGAGCGAGCTGAAGCCCGCGCCGCGACGAA 61
 61 ACTTCGG 120
 62 ACTTCGG 121
 121 AGCCCGGACCGCGTCAAGCGGCCCATGAATGCCCTTCATGCTGTGTCCTCCGCGGACGCG 180
 122 AGCCCGGACCGCGTCAAGCGGCCCATGAATGCCCTTCATGCTGTGTCCTCCGCGGACGCG 181
 181 CGCAAGATGGCCCGAGAGAAACCCAGATGCACTCGGAGATCGCAAGGCGCTGGGCG 240
 182 CGCAAGATGGCCCGAGAGAAACCCAGATGCACTCGGAGATCGCAAGGCGCTGGGCG 241
 241 GCGGAGTGAACCTTTGTTCGAGAGAGAGAGCGGCGCTTCATCGACGAGGCTTAAGCGG 300
 242 GCGGAGTGAACCTTTGTTCGAGAGAGAGAGCGGCGCTTCATCGACGAGGCTTAAGCGG 301
 301 CTGCGAGCGCTGCACTGAAGAGAGAGAGCGGCGGATTAATAATACCGCGCGCGGAAAC 360
 302 CTGCGAGCGCTGCACTGAAGAGAGAGAGCGGCGGATTAATAATACCGCGCGCGGAAAC 361
 361 AAGACGCTCATGAAGAAGATTAAGTACAGCTGCCGCGGCGGCTGTGCGCCCGCGGCGG 420
 421 AATAGCATGGCGAGGG 480
 422 AATAGCATGGCGAGCGGNGTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 481
 481 ATGGACAGTTAGCGGCACAT 500
 482 ATGGACAGTTAGCGGCACAT 501
 RESULT 7

BE206866
 LOCUS 641 bp mRNA linear EST 27-JUN-2000
 DEFINITION ba04909.y1 NIH-MGC.7 Homo sapiens cDNA clone IMAGE:2823424 5', similar to SW:SOX2_HUMAN P48431 TRANSCRIPTION FACTOR SOX-2. [1] ;, mRNA sequence.
 ACCESSION BE206866
 VERSION BE206866.1 GI:8750264
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 641)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/btml/tiresources.shtml
 Seq primer: -400P from Gibco
 High quality sequence stop: 518.
 Location/Qualifiers

FEATURES

source
 1. 641
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2823424"
 /clone_lib="NIH-MGC.7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pOTB1; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(5). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 136 a 205 c 231 g 69 t
 ORIGIN

Query Match 99.4%; Score 496.8; DB 9; Length 641;
 Best Local Similarity 99.6%; Pred. No. 2.9e-80;
 Matches 498; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 CACAGCGCCCGCATGTATACATGATGAGAGCGAGCTGAAGCCCGCGCGCGACGAA 60
 63 CACAGCGCCCGCATGTATACATGATGAGAGCGAGCTGAAGCCCGCGCGCGACGAA 122
 61 ACTTCGG 120
 123 ACTTCGG 182
 121 AGCCCGGACCGCGTCAAGCGGCCCATGAATGCCCTTCATGCTGTGTCCTCCGCGGACGCG 180
 183 AGCCCGGACCGCGTCAAGCGGCCCATGAATGCCCTTCATGCTGTGTCCTCCGCGGACGCG 242
 181 CGCAAGATGGCCCGAGAGAAACCCAGATGCACTCGGAGATCGCAAGGCGCTGGGCG 240
 243 CGCAAGATGGCCCGAGAGAAACCCAGATGCACTCGGAGATCGCAAGGCGCTGGGCG 302
 241 GCGGAGTGAACCTTTGTTCGAGAGAGAGAGCGGCGGATTAATAATACCGCGCGCGGAAAC 300
 303 GCGGAGTGAACCTTTGTTCGAGAGAGAGAGCGGCGGATTAATAATACCGCGCGCGGAAAC 362
 301 CTGCGAGCGCTGCACTGAAGAGAGAGCGGCGGATTAATAATACCGCGCGCGGAAAC 360
 363 CTGCGAGCGCTGCACTGAAGAGAGAGCGGCGGATTAATAATACCGCGCGCGGAAAC 422

QY 361 AAGACCTCATGAGAGATTAAGTACACGCTGCCCCGCGGCTGCTGGCCCCCGCGGC 420
 DB 423 AAGACCTCATGAGAGATTAAGTACACGCTGCCCCGCGGCTGCTGGCCCCCGCGGC 482
 QY 421 AATGATGATGCGAGCGGCGGCTGCGGCGCGGCTGCGGCGGCGGCTGGAACGAGCG 480
 DB 483 AATGATGATGCGAGCGGCGGCTGCGGCGCGGCTGCGGCGGCGGCTGGAACGAGCG 542
 QY 481 ATGACAGTTACGCGACAT 500
 DB 543 ATGACAGTTACGCGACAT 562

RESULT 8
 LOCUS BE672418 728 bp mRNA linear EST 08-SEP-2000
 DEFINITION BE672418.1 NCI-CGAP GC6 Homo sapiens cDNA clone IMAGE:3223136 3' similar to SW:SOX2_HUMAN P48431 TRANSCRIPTION FACTOR SOX-2. [1]
 ; contains TARI.T1 TARI repetitive element; mRNA sequence.

ACCESSION BE672418
 VERSION BE672418.1 GI:10032959
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 728)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D., M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seg primer: -400p from Gibco
 High quality sequence stop: 497.
 Location/Qualifiers
 1..728

FEATURES
 SOURCE
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3223136"
 /clone_id="NCI-CGAP GC6"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 162 a 236 c 245 g 84 t
 ORIGIN

Query Match 99.4%; Score 496.8; DB 10; Length 728;
 Best Local Similarity 99.6%; Pred. No. 2.9e-80;
 Matches 498; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CACAGCGCCCGATGACACATGATGAGAGCGAGCTAAGCCGCGGCGCCGACGAA 60
 DB 27 CACAGCGCCCGATGACACATGATGAGAGCGAGCTAAGCCGCGGCGCCGACGAA 86

QY 61 ACTTGGGGGGGGGCGGCGGCGGCACTCCACCGGCGGCGGCGGCGGCAACCGAAGAAAC 120
 DB 87 ACTTGGGGGGGGGCGGCGGCGGCAACTCCACCGGCGGCGGCGGCGGCAACCGAAGAAAC 146
 QY 121 AGCCCGAGACCGGCTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 147 AGCCCGAGACCGGCTCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 206
 QY 181 CGCAGATGCGCCGAGAGAACCCCAAGATGCANACTGGAGATATAGCAACGCTGGGC 240
 DB 207 CGCAGATGCGCCGAGAGAACCCCAAGATGCANACTGGAGATATAGCAACGCTGGGC 266
 QY 241 GCCGAGTGGAAACTTCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 DB 267 GCCGAGTGGAAACTTCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326
 QY 301 CTGGGAGGCTGCACATGAG 360
 DB 327 CTGGGAGGCTGCACATGAG 386
 QY 361 AAGACCTCATGAGAGATTAAGTACACGCTGCCCCGCGGCTGCTGGCCCCCGCGGC 420
 DB 387 AAGACCTCATGAGAGATTAAGTACACGCTGCCCCGCGGCTGCTGGCCCCCGCGGC 446
 QY 421 AATGATGATGCGAGCGGCGGCTGCGGCGCGGCTGCGGCGGCGGCTGGAACGAGCG 480
 DB 447 AATGATGATGCGAGCGGCGGCTGCGGCGCGGCTGCGGCGGCGGCTGGAACGAGCG 506
 QY 481 ATGACAGTTACGCGACAT 500
 DB 507 ATGACAGTTACGCGACAT 526

RESULT 9
 LOCUS A1571299 507 bp mRNA linear EST 12-MAY-1999
 DEFINITION A1571299.tn44e04.x1 NCI-CGAP Brn25 Homo sapiens cDNA clone IMAGE:2170494 3' similar to gb:U11136 SOX-10 PROTEIN (HUMAN); contains TARI.T2 TARI repetitive element; mRNA sequence.

ACCESSION A1571299
 VERSION A1571299.1 GI:4534673
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 507)
 NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/Brn25), Tumor Gene Index
 JOURNAL Unpublished (1998)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrp/image/image.html
 Insert length: 1184 Std Error: 0.00
 Seg primer: -400p from Gibco
 High quality sequence stop: 451
 POLYA=No.
 Location/Qualifiers
 1..507

FEATURES
 SOURCE
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2170494"
 /clone_id="NCI-CGAP_Brn25"

Db 481 ATGCAGATTACGCGCACAT 500

RESULT 11
LOCUS A1356682 965 bp mRNA linear EST 15-FEB-1999
DEFINITION g122a10.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:2012730 3'
similar to gb:X71136 SOX-10 PROTEIN (HUMAN);, mRNA sequence.

ACCESSION A1356682
VERSION A1356682.1 GI:4108303
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 965)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bhrp/image/image.html
Insert Length: 1190 Std Error: 0.00
Seq primer: -400p from G1bco
High quality sequence stop: 480.
Location/Qualifiers
1. 965
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2012730"
/clone_1lb="NCI CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCATATCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldi."

BASE COUNT 202 a 329 c 308 g 125 t 1 others

ORIGIN

Query Match 97.6%; Score 488; DB 9; Length 965;
Best Local Similarity 99.8%; Pred. No. 1.2e-78;
Matches 499; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CACAGCGCCGATGTACATCATGTGAGAGAGAGTGAAGCCGCGCGCCGACGAA 60
|||||
Db 27 CACAGCGCCGATGTACATCATGTGAGAGAGTGAAGCCGCGCGCCGACGAA 86
|||||

QY 61 ACTTCGGGGGGGGGGGGGACACTCCACCGCGGGCGGGCGGCGCAACAGAAAAA 120
|||||
Db 87 ACTTCGGGGGGGGGGGGGACACTCCACCGCGGGCGGGCGGCGCAACAGAAAAA 146
|||||

QY 121 AGCCGGGACCGGTCACAGCGCCCATGTAATGCTTTCATGCTGTGCTCCGCGGACAGGG 180
|||||
Db 147 AGCCGGGACCGGTCACAGCGCCCATGTAATGCTTTCATGCTGTGCTCCGCGGACAGGG 206
|||||

QY 181 CGCAAGATGGCCCGAGAGAACCCCAAGATGCACAACCTGGAGATCAGCAACCGCTGGGC 240
|||||

Db 207 CGCAAGATGGCCCGAGAGAACCCCAAGATGCACAACCTGGAGATCAGCAACCGCTGGGC 266

QY 241 GCCAGTGTGAACCTTTTGTTCGAGACGAGAGACCGGCGCTTCATGACGAGGCTAAGCG 300
|||||

Db 267 GCCAGTGTGAACCTTTTGTTCGAGACGAGAGACCGGCGCTTCATGACGAGGCTAAGCG 326
|||||

QY 301 CTGGAGCGGTGACATGAGAGAGACCGCGGATTTATTAATTCGCGCGCGGAGAAAC 360
|||||

Db 327 CTGGAGCGGTGACATGAGAGAGACCGCGGATTTATTAATTCGCGCGCGGAGAAAC 386
|||||

QY 361 AAGACGCTCATGAGAGAGATTAATGATACGCTCCCGCGGCTGCTGCGCCCGCGCG 420
|||||

Db 387 AAGACGCTCATGAGAGAGATTAATGATACGCTCCCGCGGCTGCTGCGCCCGCGCG 445
|||||

QY 421 AATAGCATGCGAGCGGGGTGCGGGGTGCGGGCGCGGCTGCGGGCGGAGCAACAGCG 480
|||||

Db 446 AATAGCATGCGAGCGGGGTGCGGGGTGCGGGCGCGGCTGCGGGCGGAGCAACAGCG 505
|||||

QY 481 ATGCAGATTACGCGCACAT 500
|||||

Db 506 ATGCAGATTACGCGCACAT 525
|||||

RESULT 12
LOCUS AM055151 607 bp mRNA linear EST 09-MAR-2000
DEFINITION w201d08.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:2556783 3'
similar to gb:X71136 SOX-10 PROTEIN (HUMAN);, mRNA sequence.

ACCESSION AM055151
VERSION AM055151.1 GI:5920854
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 607)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bhrp/image/image.html
Insert Length: 1252 Std Error: 0.00
Seq primer: -400p from G1bco
High quality sequence stop: 501.
Location/Qualifiers
1. 607
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2556783"
/clone_1lb="NCI CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCATATCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldi."

BASE COUNT 133 a 192 c 214 g 64 t 4 others

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Insert length: 1142 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 485
 POLYA-No.

FEATURES

Source 1..514 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2214163"
 /clone_1ib="NCI-CGAP.Lu19"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 -R1 adaptor primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

BASE COUNT 110 a 160 c 190 g 54 t

ORIGIN

Query Match 95.4%; Score 477; DB 9; Length 514;
 Best Local Similarity 99.8%; Pred. No. 1e-76;
 Matches 488; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 CACAGCGCCCGCATGTACACATGATGAGAGCGAGCTGAAGCCCGCGGCGCCAGCA 60
 |||||||
 27 CACAGCGCCCGCATGTACACATGATGAGAGCGAGCTGAAGCCCGCGGCGCCAGCA 86
 |||||||
 61 ACTTCGGGGGGCGCGCGCACTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 |||||||
 87 ACTTCGGGGGGCGCGCGCGCACTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 146
 |||||||
 121 AGCCCGAGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 |||||||
 147 AGCCCGAGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 206
 |||||||
 181 CGCAAGATGGCCAGAGAGACCCCAAGATGACAACTCGAGATCGAAGCCCTGGCG 240
 |||||||
 207 CGCAAGATGGCCAGAGAGACCCCAAGATGACAACTCGAGATCGAAGCCCTGGCG 266
 |||||||
 241 GCCGATGGAACCTTTTGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 |||||||
 267 GCCGATGGAACCTTTTGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326
 |||||||
 301 CTCGAGCGCTGCACATGAAGAGACCGCGGATTTAATATACCGCGCGCGCGGAGAAAC 360
 |||||||
 327 CTCGAGCGCTGCACATGAAGAGACCGCGGATTTAATATACCGCGCGCGCGGAGAAAC 386
 |||||||
 361 AAGAGCGCTATGAAGAGAGATAGTACAGCTGCGCGCGCGCGCTGCTGGCGCGCGCGCG 420
 |||||||
 387 AAGAGCGCTATGAAGAGAGATAGTACAGCTGCGCGCGCGCGCTGCTGGCGCGCGCGCG 446
 |||||||

QY 421 AATAGCATGGCGAGCGGGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 |||||||
 Db 447 AATAGCATGGCGAGCGGGGCTC-66GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 505
 |||||||
 QY 481 ATGAGCAGT 489
 |||||||
 Db 506 ATGAGCAGT 514

RESULT 15
 BE792833 740 bp mRNA linear EST 20-SEP-2000
 BE792833
 LOCUS 601584834F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3939033 5',
 mRNA sequence.
 DEFINITION BE792833
 ACCESSION BE792833
 VERSION BE792833.1 GI:10214031
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 740)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE NIH-MGC <http://imgc.ncbi.nlm.nih.gov/>.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LICM787 row: d column: 10
 High quality sequence stop: 740.

FEATURES

Source 1..740 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3939033"
 /clone_1ib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pOB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCCACGAG(C). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 158 a 240 c 249 g 93 t

ORIGIN

Query Match 93.4%; Score 467; DB 10; Length 740;
 Best Local Similarity 100.0%; Pred. No. 6.8e-75;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAGCTGAAGCG 93
 |||||||
 Db 1 GAGCTGAAGCG 60
 |||||||
 QY 94 GCGGCG 153
 |||||||
 Db 61 GCGGCG 120
 |||||||
 QY 154 TTCATGCTGTGTCG 213
 |||||||
 Db 121 TTCATGCTGTGTCG 180
 |||||||
 QY 214 AACTCGAGATGAGCAAGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
 |||||||

```

Db 181 AACTGGAGATCAGCAAGCCCTGGGGCCCGAGTGAACCTTTGTGAGAGACGAGAAAG 240
QY 274 CGGCCGTTTCATCGACGAGGCTAAAGCGGCTGCGAGCGCTGCACATGAAGAGACCCGGAT 333
    |||||||
Db 241 CGGCCGTTTCATCGACGAGGCTAAAGCGGCTGCGAGCGCTGCACATGAAGAGACCCGGAT 300
QY 334 TATTAATACCGGCCCCCGGCGAAACCAAGACGCTTCATGAAGAAGATTAAGTACAGCGTG 393
    |||||||
Db 301 TATTAATACCGGCCCCCGGCGAAACCAAGACGCTTCATGAAGAAGATTAAGTACAGCGTG 360
QY 394 CCGGCGGGGCTGCTGGCCCCCGGCGCAATAGCATGCGAGCGGGTGGGGTGGGGGCC 453
    |||||||
Db 361 CCGGCGGGGCTGCTGGCCCCCGGCGCAATAGCATGCGAGCGGGTGGGGTGGGGGCC 420
QY 454 GGCTGGGGCGGGGCGTGAACCAAGGCAATGACAGTTACGGGCACAT 500
    |||||||
Db 421 GGCTGGGGCGGGGCGTGAACCAAGGCAATGACAGTTACGGGCACAT 467

```

Search completed: October 10, 2002, 20:27:13
 Job time : 702.733 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 10, 2002, 12:39:37 ; Search time 84.9333 Seconds

(without alignments)
10107.424 Million cell updates/sec

Title: US-09-489-101a-3_COPY_1_500

Perfect score: 500
Sequence: 1 cacagcgccgcagctacacaa.....atgacagcttacgcgcacat 500

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*

```
1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	1085	22	AAD11110 Human small cell 1
2	500	100.0	1087	22	AAK52388 Human polynucleoti
3	500	100.0	1116	22	AAK53268 Human polynucleoti
4	500	100.0	1126	22	AAK53284 Human polynucleoti
5	500	100.0	1501	22	AAK53372 Human polynucleoti
6	432.8	86.6	2283	24	AB199318 Mouse ischaemic co
7	350	70.0	395	24	AAK51666 Lung small cell ca
8	279	55.8	1542	20	AAK16151 Human Sox1 encodin
9	279	55.8	4091	22	AAD11111 Human small cell 1

10	266.2	53.2	2312	20	AAK16152 Chicken Sox1 cDNA.
11	261.4	52.0	2376	20	AAK16153 Mouse Sox1 cDNA.
12	234.8	47.0	2309	22	AAD11118 Human small cell 1
13	225	45.0	8372	22	AAD11119 Human small cell 1
14	220	44.0	2378	23	AAK87738 DNA encoding novel
15	217.8	43.6	412	24	AAK51625 Lung small cell ca
16	209.2	41.8	1161	20	AAK16154 Degenerate human S
17	199	39.8	1969	23	AAK87779 DNA encoding novel
18	193.4	38.7	2298	23	AB18481 Drosophila melanog
19	193.4	38.7	4298	23	AB18480 Drosophila melanog
20	150.8	30.2	1788	22	AAK16401 Human cDNA sequenc
21	143.2	28.6	3934	17	AAAT30309 Human SOX-9 cDNA.
22	140.8	28.2	434	23	AAK87736 DNA encoding novel
23	140	28.0	2514	17	AAAT30308 Human gastric can
24	140	28.0	2514	17	AAAT30308 Mouse Sox-9 cDNA.
25	138.4	27.7	2514	13	AAK20684 P53.3 (NCIMB 4030
26	134.6	26.9	1661	16	AAK24814 Bovine sex-determi
27	134.6	26.9	1661	16	AAK24811 Bovine sex-determi
28	134.6	26.9	1661	16	AAK24813 Sex-determining re
29	134.6	26.9	2840	17	AAAT33008 Rabbit mt-box. Or
30	134.2	26.8	996	13	AAK20680 SOX4 coding sequen
31	128.6	25.7	1422	18	AAAT59844 Porcine sex-determ
32	128.2	25.6	1667	16	AAK24812 Mouse ischaemic co
33	127.4	25.5	2069	24	AB199491 Lung small cell ca
34	121.4	24.3	441	24	AAK51805 Clone WGT24 of a g
35	120.6	24.1	2033	21	AAK61724 Mouse ischaemic co
36	119	23.8	2215	24	AB199658 Drosophila melanog
37	117.6	23.5	2310	23	AB105611 Mouse mt-box. Mus
38	117.4	23.5	471	13	AAK20681 Mouse SRY-related
39	117.4	23.5	10266	17	AAAT33007 PK5 741 insert con
40	117.4	23.5	14704	13	AAK20685 Human cervical can
41	117.2	23.4	4467	22	AAAT2650 Drosophila melanog
42	115.2	23.0	7108	23	AB105610 Fish protamine gen
43	112.6	22.5	2762	16	AAK27042 Mouse Sox-18 gene.
44	106.2	21.2	1588	18	AAK16167 Mouse Sox-18 gene.
45	103.8	20.8	317	16	AAAT3968 Bovine testicle de

ALIGNMENTS

RESULT 1	
AAD11110	standard; DNA; 1085 BP.
ID	AAD11110;
AC	24-SEP-2001 (first entry)
XX	
DE	Human small cell lung cancer associated gene, SOX2.
XX	
XX	Human; small cell lung cancer; therapy: hCAAP; nucleic acid; NA;
KW	melanoma; cancer; colon; breast; head; neck; transitional cancer;
KW	leiomyosarcoma; synovial sarcoma; cytostatic; SOX2; ds.
XX	
OS	Homo sapiens.
XX	
FT	key
FT	Location/Qualifiers
FT	CDS
FT	13..966
FT	/*tag= a
FT	/product= "Human SOX2 protein"
XX	
XX	
XX	
PD	26-JUL-2001.
PF	19-JAN-2001; 2001WO-US02015.
XX	
PR	21-JAN-2000; 2000US-0489101.
XX	
PA	(LUDW-) LUDWIG INST. CANCER RES.
PA	(SLOK) SLOAN KETTERING INST. CANCER RES.
PA	(CORR) CORNELL RES. FOUND INC.

Db 154 AGCCGGACCCGCTCAAGCGGGCCCATGATGCTTTCATGCTGTGTCTCCGGGGACACCG 213
 Oy 181 CGAAGATGCGCCAGAGAGACCCCAAGATGACAACTCGGAGATCAGCAAGCGCTGGGC 240
 Db 214 CGCAAGATGCGCCAGAGAGACCCCAAGATGACAACTCGGAGATCAGCAAGCGCTGGGC 273
 Oy 241 GCCGAGTGAAGCTTTTGTGAGAGCGAGAAAGCGGCGTTTCATCGAGCAAGCTTAAGCGG 300
 Db 274 GCCGAGTGAAGCTTTTGTGAGAGCGAGAAAGCGGCGTTTCATCGAGCAAGCTTAAGCGG 333
 Oy 301 CTGCGAGCGCTGACATGAAAGAGACCCGAGTTTAAATACCGGCCCGCGGAAAAAC 360
 Db 334 CTGCGAGCGCTGACATGAAAGAGACCCGAGTTTAAATACCGGCCCGCGGAAAAAC 393
 Oy 361 AAGACGCTCATGAAGAAAGATTAAGTACAGCTGCGCGCGGGCGCTGCTGGCGCGCGGCG 420
 Db 394 AAGACGCTCATGAAGAAAGATTAAGTACAGCTGCGCGCGGGCGCTGCTGGCGCGCGGCG 453
 Oy 421 AATAGCATGCGAGCGGGGTGCGGGTGGCGCCCGGCTGGCGCGCGGCGTGAACGCGC 480
 Db 454 AATAGCATGCGAGCGGGGTGCGGGTGGCGCCCGGCTGGCGCGCGGCGTGAACGCGC 513
 Oy 481 ATGACAGTTACGCGACAT 500
 Db 514 ATGACAGTTACGCGACAT 533

RESULT 3

ID AAK53268 standard; cDNA; 1116 BP.
 AAK53268;

DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 2797.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 PN WO200157190-A2.
 PD 09-AUG-2001.
 PP 05-FEB-2001; 2001WO-US04098.
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 PI Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZH;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM80135.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PS useful in diagnosis and gene therapy -
 PS Claim 1; Page 4995; 6221pp; English.

CC The invention relates to polynucleotides (AAK51456-AAK51435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibit activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM60020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 1116 BP; 256 A; 362 C; 362 G; 136 T; 0 other:

Query Match 100.0%; Score 500; DB 22; Length 1116;
 Best Local Similarity 100.0%; Pred. No. 7.7e-95;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CACAGCGCCGCGATGTAACAATGATGAGAGACGAGCTGAAGCCGCGCGCAAGCAA 60
 Db 27 CACAGCGCCGCGATGTAACAATGATGAGAGACGAGCTGAAGCCGCGCGCAAGCAA 86
 Oy 61 ACTTGG 120
 Db 87 ACTTGG 146
 Oy 121 AGCCCGAGCCGCTCAAGCGGGCCCATGAATGCTTATGTTGTGTCTCCGCGGCGAGCGG 180
 Db 147 AGCCCGAGCCGCTCAAGCGGGCCCATGAATGCTTATGTTGTGTCTCCGCGGCGAGCGG 206
 Oy 181 CGCAAGATGCGCCAGAGAGAAACCCCAAGATGACAACTCGGAGATGAGAGAGCGCTGGGC 240
 Db 207 CGCAAGATGCGCCAGAGAGAAACCCCAAGATGACAACTCGGAGATGAGAGAGCGCTGGGC 266
 Oy 241 GCCGAGTGAAGCTTTTGTGAGAGCGAGAAAGCGGCGTTTCATCGAGCAAGCTTAAGCGG 300
 Db 267 GCCGAGTGAAGCTTTTGTGAGAGCGAGAAAGCGGCGTTTCATCGAGCAAGCTTAAGCGG 326
 Oy 301 CTGCGAGCGCTGACATGAAAGAGACCCGAGTTTAAATACCGGCCCGCGGAAAAAC 360
 Db 327 CTGCGAGCGCTGACATGAAAGAGACCCGAGTTTAAATACCGGCCCGCGGAAAAAC 386
 Oy 361 AAGACGCTCATGAAGAAAGATTAAGTACAGCTGCGCGCGGGCGTGGCGCGCGCGCGCG 420
 Db 387 AAGACGCTCATGAAGAAAGATTAAGTACAGCTGCGCGCGGGCGTGGCGCGCGCGCGCG 446
 Oy 421 AATAGCATGCGAGCGGGGTGCGGGTGGCGCCCGGCTGGCGCGCGGCGTGAACGCGC 480
 Db 447 AATAGCATGCGAGCGGGGTGCGGGTGGCGCCCGGCTGGCGCGCGGCGTGAACGCGC 506
 Oy 481 ATGACAGTTACGCGACAT 500
 Db 507 ATGACAGTTACGCGACAT 526

RESULT 4

ID AAK52284 standard; cDNA; 1126 BP.
 AAK52284;

AC 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 829.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.
 XX MO200157190-A2.
 XX 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001MO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AAM79151.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PS Claim 1; Page 2766-2767; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK60302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM6020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX Sequence 1126 BP; 262 A; 361 C; 364 G; 139 T; 0 other;
 SQ
 Query Match 100.0%; Score 500; DB 22; Length 1126;
 Best Local Similarity 100.0%; Pred. No. 7,7e-95;
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CACAGCGCGCGCATGTACACATATGAGAGAGCTGAAGCCCGCGGCGCCGACGAA 60
 DB 34 CACAGCGCGCGCATGTACACATATGAGAGAGCTGAAGCCCGCGGCGCCGACGAA 93
 QY 61 ACTTCGGGGGGGGGGCGGCGGCACTCCACCGGGCGGCGGCGCAACCAAGAAAAAC 120
 DB 94 ACTTCGGGGGGGGGGCGGCGGCACTCCACCGGGCGGCGGCGGCAACCAAGAAAAAC 153
 QY 121 AGCCCGGAGCCCGCTGACAGCGGCGCATGATGCTTCAATGGTGTGTCCCGGGGACCGG 180
 DB 154 AGCCCGGAGCCCGCTGACAGCGGCGCATGATGCTTCAATGGTGTGTCCCGGGGACCGG 213
 QY 181 CGCAAGATGGCGCGGAGAACCCCAAGATGACACAACTCGAGATCAGCAAGCGCTGGGC 240
 DB 214 CGCAAGATGGCGCGGAGAACCCCAAGATGACACAACTCGAGATCAGCAAGCGCTGGGC 273
 QY 241 GCCGAGTGGAACTTTTGTGTGGAGACGAGAAAGCGGCTTCATCAGACGAGGCTAAAGCG 300
 DB 274 GCCGAGTGGAACTTTTGTGTGGAGACGAGAAAGCGGCTTCATCAGACGAGGCTAAAGCG 333
 QY 301 CTGCGAGCGCTGCACATGAAGAGACACCGCGATTATAAATACCGCGCCCGGGAACCC 360

DB 334 CTGCGAGCGCTGCACATGAAGAGACACCGGATTATAAATACCGCGCCCGGGAACCC 393
 QY 361 AAGACCTCATGAAGAAGATAGTACAGCTGCCCGCGGGGCTGTGGCCCCCGCGGC 420
 DB 394 AAGACCTCATGAAGAAGATAGTACAGCTGCCCGCGGGGCTGTGGCCCCCGCGGC 453
 QY 421 AATAGCATGGCGAGCGGGGTGGGGTGGCGCGCTGGGGCGGGCGCTGAACACGCGC 480
 DB 454 AATAGCATGGCGAGCGGGGTGGGGTGGCGCGCTGGGGCGGGCGCTGAACACGCGC 513
 QY 481 ATGACAGTTACGCGACAT 500
 DB 514 ATGACAGTTACGCGACAT 533
 RESULT 5
 AAK53372
 ID AAK53372 standard; cDNA; 1501 BP.
 XX
 AC AAK53372;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 2901.
 XX
 KW Human: cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 XX MO200157190-A2.
 XX 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001MO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AAM60239.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PS Claim 1; Page 5070; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK60302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAB60020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 1501 BP; 391 A; 428 C; 450 G; 220 T; 12 other;

Query Match 100.0%; Score 500; DB 22; Length 1501;

Best Local Similarity 100.0%; Pred. No. 7.8e-95;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CACAGCGCCCGCATGACATGATGAGAGCGAGCTGAACCGCGCGCGCGAGCAA 60
   |||||||
DB 27 CACAGCGCCCGCATGACATGATGAGAGCGAGCTGAACCGCGCGCGCGAGCAA 86
   |||||||
QY 61 ACTTCGGGGGGGGGGGGGGGCAACTCCACCGCGGGCGCGCGCGCAACGAAAAAC 120
   |||||||
DB 87 ACTTCGGGGGGGGGGGGGGGCAACTCCACCGCGGGCGCGCGCGCAACGAAAAAC 146
   |||||||
QY 121 AGCCCGGACCGCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
   |||||||
DB 147 AGCCCGGACCGCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 206
   |||||||
QY 181 CGCAGAGTGGCCGAGAGAACCCCAAGATGACACACTCGAGATCAGAACGCTGGGC 240
   |||||||
DB 207 CGCAGAGTGGCCGAGAGAACCCCAAGATGACACACTCGAGATCAGAACGCTGGGC 266
   |||||||
QY 241 GCCAGTGGAACTTTTGTGCGAGAGCGAGAGCGCGCGCTTCATGACGAGCTAAAGCG 300
   |||||||
DB 267 GCCAGTGGAACTTTTGTGCGAGAGCGAGAGCGCGCGCTTCATGACGAGCTAAAGCG 326
   |||||||
QY 301 CTGCGAGCGCTGCACATGAAGAGCAACCGGATTAATAATACCGCGCGCGCGGAAAAAC 360
   |||||||
DB 327 CTGCGAGCGCTGCACATGAAGAGCAACCGGATTAATAATACCGCGCGCGCGGAAAAAC 386
   |||||||
QY 361 AAGAGCTCATGAAGAAGATGAATGACAGCTGCCCGCGCGCGCTGTGCGCGCGCGCG 420
   |||||||
DB 387 AAGAGCTCATGAAGAAGATGAATGACAGCTGCCCGCGCGCGCTGTGCGCGCGCGCG 446
   |||||||
QY 421 AATAGCATGGCGAGCGGGTGGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
   |||||||
DB 447 AATAGCATGGCGAGCGGGTGGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 506
   |||||||
QY 481 AATGACAGTTACGCGCACAT 500
   |||||||
DB 507 AATGACAGTTACGCGCACAT 526
   |||||||

RESULT 6
AB199318
ID AB199318 standard; cDNA; 2283 BP.
XX
AC AB199318;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischemic condition related cDNA sequence SEQ ID NO:179.
XX
KW Mouse; ischemia; compressive ischemia; occlusive ischemia;
XX vaspasptic ischemia; ischemic condition; ischemic disease; ss.
XX
OS Mus musculus.
XX
PN MO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001MO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UNNT-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
```

XX WPI; 2002-034733/04.
DR P-PSDB; ABB57082.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
PS Claim 2; Page 518-521; 2690pp; English.

The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischemia, occlusive ischemia or vaspasptic ischemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or therapeutics for ischemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX Sequence 2283 BP; 597 A; 583 C; 614 G; 488 T; 1 other;

Query Match 86.6%; Score 432.8; DB 24; Length 2283;

Best Local Similarity 92.5%; Pred. No. 6.6e-81;
Matches 468; Conservative 0; Mismatches 32; Indels 6; Gaps 1;

```
QY 1 CACAGCGCCCGCATGACATGATGAGAGCGAGCTGAACCGCGCGCGCGAGCAA 60
   |||||||
DB 350 CCGAGCGCCCGCATGATGATGAGAGCGAGCTGAACCGCGCGCGCGAGCAA 409
   |||||||
QY 61 ACTTCGGGGGGGGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 114
   |||||||
DB 410 GCTTCGGGGGGGGGGGGGCGCGCGCGAGAGCAACCGCGCGCGCGCGCGCGCG 469
   |||||||
QY 415 AAAAAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 174
   |||||||
DB 470 AAGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529
   |||||||
QY 475 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 234
   |||||||
DB 530 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 589
   |||||||
QY 235 CTGGGCGCGCGAGTGAACCTTTTGTGCGAGAGCGAGAGCGCGCGCGCGCGCG 294
   |||||||
DB 590 CTGGGCGCGCGAGTGAACCTTTTGTGCGAGAGCGAGAGCGCGCGCGCGCGCG 649
   |||||||
QY 295 AAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 354
   |||||||
DB 650 AAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 709
   |||||||
QY 355 AAAACCAAGAGCGCTGAGAGAGAGATGAATGACAGCTGCCCGCGCGCGCGCGCG 414
   |||||||
DB 710 AAAACCAAGAGCGCTGAGAGAGAGATGAATGACAGCTGCCCGCGCGCGCGCGCG 769
   |||||||
QY 415 GCGGCGATATGATGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
   |||||||
DB 770 GCGGCGAGAGCGAGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 829
   |||||||
QY 475 CAGCGCATGAGCATGACGCGCACAT 500
   |||||||
DB 830 CAGCGCATGAGCATGACGCGCACAT 855
   |||||||
```

RESULT 7
AAS61666
ID AAS61666 standard; cDNA; 395 BP.
XX

```

AC AAS61666;
XX
XX
DT 29-JAN-2002 (first entry)
XX
XX Lung small cell carcinoma antigen, cDNA #207.
XX
XX Human; cytostatic; antitumour; lung small cell cancer antigen;
XX tumour; lung cancer; ss.
XX
XX Homo sapiens.
XX
XX W0200177168-A2.
XX
XX 18-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US1859.
XX
XX 11-APR-2000; 2000US-196780P.
XX 21-JUN-2000; 2000US-213361P.
XX 01-SEP-2000; 2000US-229763P.
XX 05-SEP-2000; 2000US-230629P.
XX 14-SEP-2000; 2000US-232565P.
XX 19-DEC-2000; 2000US-257037P.
XX 08-JAN-2001; 2001US-260796P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Lodes MJ, Wang T, Mohamath R, Indirias CY;
XX
XX WPI: 2002-010896/01.
XX
XX Lung tumour polynucleotide and polypeptides useful in therapy and
XX diagnosis of cancer especially lung cancer -
XX
XX Claim 1: Page 192: 295pp: English.
XX
XX The invention relates to novel isolated lung small cell cancer antigen
XX polynucleotides (I) and polypeptides (II) used in a method of detecting
XX cancer in a patient. The method is optionally performed by
XX utilising oligonucleotides (III), where the biological sample
XX from the patient is contacted with (III), detecting the amount of
XX polynucleotide hybridised to (III) in the sample and comparing the
XX amount of polynucleotide to a predetermined cut-off value and thereby
XX determining cancer in a patient. (I), (II) or antigen-presenting cells
XX expressing (II) is useful for stimulating and/or expanding T cells
XX specific for a tumour protein. The method comprises contacting T cells
XX with one of the components under conditions to permit the stimulation
XX and/or expansion of the cells. A composition comprising (I) is useful for
XX stimulating an immune response in a patient and for inhibiting the
XX development of a cancer especially lung cancer in a patient. An
XX isolated T cell population is useful for removing tumour cells from the
XX biological sample and for inhibiting the development of cancer in a
XX patient. AAS61460-AAS61874 represent novel human lung small cell
XX cancer antigen coding sequences of the invention.
XX
XX Sequence 395 BP; 83 A; 128 C; 144 G; 40 T; 0 other;
XX
XX Query Match 70.0%; Score 350; DB 24; Length 395;
XX Best Local Similarity 100.0%; Pred. No. 8,1e-64;
XX Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CACAGGCGCCGCTGATCAACATGATGAGAGAGAGCTGAAGCGCGCGCGCGACCA 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 46 CACAGGCGCCGCTGATCAACATGATGAGAGAGAGCTGAAGCGCGCGCGCGACCA 105
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 ACTTGGGGGGCGCGCGCACTCCACCGCGCGCGCGCGCGCAACCAAGAAAAAC 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 106 ACTTGGGGGGCGCGCGCACTCCACCGCGCGCGCGCGCGCAACCAAGAAAAAC 165
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 AGCCGCGACCGCGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 166 AGCCGCGACCGCGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 225

```

```

XX
XX 181 CGCAGATGCGCCCGAGGAGAACCCCAAGATGACACTCGAGATCAGCAAGCCCTGGGC 240
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 226 CGCAGATGCGCCCGAGGAGAACCCCAAGATGACACTCGAGATCAGCAAGCCCTGGGC 285
XX
XX 241 GCGGAGTGGAAACTTTTTCGCGAGACGAGAACCGCGCGCTTCATCGACGAGGCTAAGCG 300
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 286 GCGGAGTGGAAACTTTTTCGCGAGACGAGAACCGCGCGCTTCATCGACGAGGCTAAGCG 345
XX
XX 301 CTCGAGCGCTGCATGAGAGACACCGGATTATTAATATCCGCGCGCG 350
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 346 CTCGAGCGCTGCATGAGAGACACCGGATTATTAATATCCGCGCGCG 395
XX
XX
XX RESULT 8
XX AAX16151
XX ID AAX16151 standard; cDNA to mRNA; 1542 BP.
XX
XX AAX16151;
XX
XX 22-APR-1999 (first entry)
XX
XX Human Sox1 encoding cDNA.
XX
XX Human; Sox1; neuronal stem cell gene; neuroblastic cell; cancer;
XX nervous system; neurological disorder; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_difference 1..60
XX /*tag= a
XX /note= "Specifically claimed Sox1 control sequence"
XX CDS 60..1223
XX FT /*tag= b
XX FT /product= "Sox1"
XX
XX W09000516-A2.
XX
XX 07-JAN-1999.
XX
XX 25-JUN-1998; 98MO-GB01862.
XX
XX 25-JUN-1997; 97GB-0013469.
XX
XX (MED1-) MEDICAL RES COUNCIL.
XX
XX Lovell-Badge R, Pevery LH, Smith A;
XX
XX WPI: 1999-095759/08.
XX
XX P-PSDB; AAM94461.
XX
XX Isolating neuroblastic cells from a population - by detecting the
XX expression of the Sox1 gene in the cells and sorting the cells to
XX isolate those cells expressing Sox1
XX
XX Claim 13: Page 45-47; 60pp; English.
XX
XX A method has been developed for isolating neuroblastic cells from a cell
XX population. The method comprises: (a) detecting Sox1 gene expression in
XX the cells; and (b) isolating those cells expressing Sox1. Also described
XX is a method for producing a cell committed to the neuronal lineage,
XX comprising: (a) transfecting a pluripotent stem cell with a genetic
XX construct encoding Sox1 expression; (b) culturing the stem cells to
XX differentiate into neural cells; and (c) isolating those neural cells
XX produced. The present sequence encodes human Sox1. Detection of Sox1
XX expressing cells is important in diagnosing and treating cancers of the
XX nervous system. Neural stem cells are useful for the treatment of
XX neurological disorders, especially for repair of accidentally induced
XX trauma in the CNS or for correction of congenital or pathological
XX diseases of the CNS. A patient with a neurological disorder can act as a
XX self-donor. Cells may be isolated from the patient and either sorted to
XX extract neuroblasts, or treated in order to differentiate neuroblasts,
XX from specific or general precursors.

```

```
XX Sequence 1542 BP; 238 A; 565 C; 560 G; 179 T; 0 other:
SQ
Query Match          55.8%; Score 279; DB 20; Length 1542;
Best Local Similarity 80.3%; Pred. No. 4,5e-49;
Matches 327; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 65 CGGGGGGCGGGCGGCACTCCACCGCGGGCGCGGCAACAGAAAACAGCC 124
    ||||| ||||| ||| | | | | | | | | | | | | | | | | | | |
DB 142 CGGGGGGCGGGGGGCGGAGCGGGGCGCGGGCGGGGGCGCCCAAGGCCAAC 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 125 CGGACCGGCTCAACCGGCCCTGTAATGCTTCAATGCTGTCGCCGGCGGCGCA 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 202 AGGACCGGGTCAACCGGCCCTTCAATGCTGTCGCCGGCGGCGGCGCA 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 185 AGATGGCCCAAGAGACACCCCAAGATGCACACTCGAGATCAGCAAGCCCTGGGCGCG 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 262 AGATGGCCCAAGAGACACCCCAAGATGCACACTCGAGATCAGCAAGCCCTGGGCGCG 321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 245 AGTGAACCTTTGTCTCGGAGACGAGAAAGCGCGCTTCATCGACGAGCTAAGCGGCTGC 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 322 AGTGAACCTTTGTCTCGGAGACGAGAAAGCGCGCTTCATCGACGAGCTAAGCGGCTGC 381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 305 GAGGCGTCCACATGAGAGACACCCGATTAATAATACCGCGCGCGGCAAAACCAAGA 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 382 GCGGCTGCACATGAGAGACACCCGATTAATAATACCGCGCGCGGCAAAACCAAGA 441
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 365 CGCTCATGAAGAGATTAAGTACACGCTGCCGCGGGGCTGCTGGCCCGCGGCAATA 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 442 CGCTCATGAAGAGATTAAGTACACGCTGCCGCGGGGCTGCTGGCCCGCGGCAATA 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 425 GCATGGCGAGCGGGGTGCGGGTGGGCGCGCGCTTGGCGCGGCGCGTGC 471
    || | | | | | | | | | | | | | | | | | | | | | | | |
DB 502 GCGGCGCGCGGCTGTGGCCATGGGCGTGGGCGTGGCGCTGGGCGCG 548
    || | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
AADI1111
ID AADI1111 standard; DNA; 4091 BP.
AC AADI1111;
XX
XX 24-SEP-2001 (first entry)
DE Human small cell lung cancer associated gene, SOX1.
XX
XX Human; small cell lung cancer; therapy; hCAAP; nucleic acid; NA;
KW melanoma; cancer; colon; breast; head; neck; transitional cancer;
KW leiomyosarcoma; synovial sarcoma; cytostatic; SOX1; ds.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 61..1224
FT /*tag= a
FT /product= "Human SOX1 protein"
XX
XX WO200153349-A2.
FN
XX
XX 26-JUL-2001.
PD
XX
XX 19-JAN-2001; 2001WO-US02015.
PF
XX
XX 21-JAN-2000; 2000US-0489101.
PR
XX
XX (LUDWIG) LUDWIG INST CANCER RES.
PA (SLOAN KETTERING INST CANCER RES.
PA (CORR ) CORNELL RES FOUND INC.
XX
XX Stockert E, Scanlan MJ, Jager D, Old LJ, Gure AO, Chen Y;
XX WPI; 2001-457597/49.
XX
XX P-PSDB; AAE05811.
XX
```

```
XX Isolated polypeptide, used to treat or prognose a disorder
PT characterized by expression of a hCAAP e.g. cancer, is encoded by an
PS isolated nucleic acid comprising an NA Group 3 or 4 molecule -
XX
XX Claim 57; Page 88-90; 152pp; English.
XX
XX The invention relates to nucleic acids and encoded polypeptides which
CC are cancer associated antigens expressed in patients afflicted with
CC small cell lung cancer. The molecules provided by the invention can be
CC used in the diagnosis, monitoring, research or treatment of conditions
CC characterised by the expression of one or more cancer associated
CC antigens. The polypeptide is used to treat a disorder characterised by
CC expression of a hCAAP, and determine regression, progression or onset
CC of a condition characterised by expression of an abnormal amount of a
CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
CC are small and non-small cell lung cancer, melanoma, colon, breast, head
CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
CC The present sequence is a small cell lung cancer associated gene
CC designated as NT-SCLC-2, encoding human SOX1 protein.
XX
XX Sequence 4091 BP; 924 A; 1109 C; 1110 G; 943 T; 5 other:
SQ
Query Match          55.8%; Score 279; DB 22; Length 4091;
Best Local Similarity 80.3%; Pred. No. 4,7e-49;
Matches 327; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 65 CGGGGGGCGGGCGGCACTCCACCGCGGGCGCGGCGGCAACCAAGAAAACAGCC 124
    ||||| ||||| ||| | | | | | | | | | | | | | | | | | | |
DB 143 CGGGGGGCGGGGGGCGGAGCGGGGCGGGCGGGCGGGGCCCAAGGCCAACCC 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 125 CGGACCGGCTCAACGGGCGCCATGATGCTTCATGCTGATGCTGCCGGGAGCGAGCGCA 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 203 AGGACCGGCTCAACGGGCGCCATGATGCTTCATGCTGATGCTGCCGGGAGCGAGCGCA 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 185 AGATGGCCCAAGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGCGCG 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 263 AGATGGCCCAAGAGAACCCCAAGATGCACAACTCGAGATCAGCAAGCGCTGGGCGCG 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 245 AGTGAACCTTTGTCTCGGAGACGAGAAAGCGCGCTTCATCGACGAGCTAAGCGGCTGC 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 323 AGTGAACCTTTGTCTCGGAGACGAGAAAGCGCGCTTCATCGACGAGCTAAGCGGCTGC 382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 305 GAGGCGTCCACATGAGAGAGACACCCGATTAATAATACCGCGCGCGGCAAAACCAAGA 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 383 GCGGCTGCACATGAGAGAGACACCCGATTAATAATACCGCGCGCGGCAAAACCAAGA 442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 365 CGCTCATGAAGAGATTAAGTACACGCTGCCGCGGGGCTGCTGGCCCGCGGCAATA 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 443 CGCTCATGAAGAGATTAAGTACACGCTGCCGCGGGGCTGCTGGCCCGCGGCAATA 502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 425 GCATGGCGAGCGGGGTGCGGGTGGGCGCGCGCTTGGCGCGGCGCGGCGTGC 471
    || | | | | | | | | | | | | | | | | | | | | | | | |
DB 503 GCGGCGCGCGGCTGTGGCCATGGGCGTGGGCGTGGGCGTGGGCGCG 549
    || | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
AAK16152
ID AAK16152 standard; CDNA to mRNA; 2312 BP.
AC AAK16152;
XX
XX 22-APR-1999 (first entry)
DE Chicken Sox1 cDNA.
XX
XX Chicken; Sox1; neuronal stem cell gene; neuroblastic cell; cancer;
KW nervous system; neurological disorder; ss.
XX
XX Gallus gallus.
XX
XX OS Gallus gallus.
XX
XX W09900516-A2.
XX
XX
```

```

PD 07-JAN-1999.
XX
XX 25-JUN-1998; 98WO-GB01862.
XX
XX 25-JUN-1997; 97GB-0013469.
XX
XX (MED1-) MEDICAL RES COUNCIL.
XX
XX Lovell-Badge R, Pevny LH, Smith A;
XX
XX WPI; 1999-095759/08.
XX
XX Isolating neuroblastic cells from a population - by detecting the
XX expression of the Sox1 gene in the cells and sorting the cells to
XX isolate those cells expressing Sox1
XX
XX Disclosure; Page 41-43; 60pp; English.
XX
XX A method has been developed for isolating neuroblastic cells from a cell
XX population. The method comprises: (a) detecting Sox1 gene expression in
XX the cells; and (b) isolating those cells expressing Sox1. Also described
XX is a method for producing a cell committed to the neuronal lineage,
XX comprising: (a) transfecting a pluripotent stem cell with a genetic
XX construct encoding Sox1 expression; (b) culturing the stem cells to
XX differentiate into neural cells; and (c) isolating those neural cells
XX produced. The present sequence represents chicken Sox1 cDNA. Detection
XX of Sox1 expressing cells is important in diagnosing and treating cancers
XX of the nervous system. Neural stem cells are useful for the treatment of
XX neurological disorders, especially for repair of accidentally induced
XX trauma in the CNS or for correction of congenital or pathological
XX diseases of the CNS. A patient with a neurological disorder can act as a
XX self-donor. Cells may be isolated from the patient and either sorted to
XX extract neuroblasts, or treated in order to differentiate neuroblasts,
XX from specific or general precursors.
XX
XX Sequence 2312 BP; 421 A; 739 C; 714 G; 438 T; 0 other;
XX
Query Match 53.2%; Score 266.2; DB 20; Length 2312;
Best Local Similarity 77.0%; Pred. No. 2e-46;
Matches 338; Conservative 0; Mismatches 98; Indels 3; Gaps 1;
OY 42 GCGGCGGCGCGCGCAAACTTCGGGGGGGGCGCGCGCACTCCACCGCGCGCGC 101
D 259 GCGCGCGCGCGCGCGCTCTCTGGGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 318
OY 102 GCGGCGCAACCAAGAAACAGCCGCGACCGCTCAAGCGCGCCCATGAATGCTTCATGGT 161
D 319 GCGGCGCGCGCGCAAGCGGGGCGAGACCGCTGAAGCGCGCCCATGAAGCGCTTCTATGGT 378
OY 162 GTGGTCCCGCGGCGAGCGCGCGCAAGATGAGCCAGAGAAACCCCAAGATGCACACTCGGA 221
D 379 GTGCTCGGGGGGCGAGCGCGCAAGATGAGCCAGAGAAATCCCAAGATGCACAACTCGGA 438
OY 222 GATAGGCAAGCGCGCGCGGCGGAGTGAACCTTTGTCGAGAGCGAGCAACCGCGCTT 281
D 439 GATAGGCAAGCGCGCTGGGCGCGAGTGAAGGTGATGTGGGCGCGAAGCGCGCTTT 498
OY 282 CATGACGAGGCTAAGGCGCTGCGAGCGCTGCACATGAAGAGCAACCGCGATTAATAA 341
D 499 CATGACGAGGCGAAGCGCGCTGCGGCGCTGCACATGAAGAGCAACCGCGATTAATAA 558
OY 342 CCGGCGCGCGGGAAGCAAGAGCGCTCATGAAGAGTAAGTAACGCTGCCGCGCG 401
D 559 CCGGCGCGCGGGAAGCAAGAGCGCTCATGAAGAGTAAGTAAGTACTGCTGCCGAGAG 618
OY 402 GCTGCTGCGCGCGCGGCAATAGCATGGGCGAGCGGGGTGCGGGGCGCGCGCGCTGGG 461
D 619 GCTGCTGCG---GCGCGCGCGCGCGCGCGCGCGCTCCCGCGTGGGCGATGGG 675
OY 462 CCGCGGCGGTGAACCAAGCGC 480
D 676 CGTGCGCGGTGATCCCGCGC 694

```

```

RESULT 11
ID AAX16153 standard; cDNA to mRNA; 2376 BP.
XX
XX AAX16153;
XX
XX 22-APR-1999 (first entry)
XX
XX Mouse Sox1 cDNA.
XX
XX Mouse; Sox1; neuronal stem cell gene; neuroblastic cell; cancer;
XX nervous system; neurological disorder; ss.
XX
XX Mus musculus.
XX
XX W09900516-A2.
XX
XX 07-JAN-1999.
XX
XX 25-JUN-1998; 98WO-GB01862.
XX
XX 25-JUN-1997; 97GB-0013469.
XX
XX (MED1-) MEDICAL RES COUNCIL.
XX
XX Lovell-Badge R, Pevny LH, Smith A;
XX
XX WPI; 1999-095759/08.
XX
XX Isolating neuroblastic cells from a population - by detecting the
XX expression of the Sox1 gene in the cells and sorting the cells to
XX isolate those cells expressing Sox1
XX
XX Disclosure; Page 43-44; 60pp; English.
XX
XX A method has been developed for isolating neuroblastic cells from a cell
XX population. The method comprises: (a) detecting Sox1 gene expression in
XX the cells; and (b) isolating those cells expressing Sox1. Also described
XX is a method for producing a cell committed to the neuronal lineage,
XX comprising: (a) transfecting a pluripotent stem cell with a genetic
XX construct encoding Sox1 expression; (b) culturing the stem cells to
XX differentiate into neural cells; and (c) isolating those neural cells
XX produced. The present sequence represents mouse Sox1 cDNA. Detection
XX of Sox1 expressing cells is important in diagnosing and treating cancers
XX of the nervous system. Neural stem cells are useful for the treatment of
XX neurological disorders, especially for repair of accidentally induced
XX trauma in the CNS or for correction of congenital or pathological
XX diseases of the CNS. A patient with a neurological disorder can act as a
XX self-donor. Cells may be isolated from the patient and either sorted to
XX extract neuroblasts, or treated in order to differentiate neuroblasts,
XX from specific or general precursors.
XX
XX Sequence 2376 BP; 439 A; 772 C; 745 G; 420 T; 0 other;
XX
Query Match 52.3%; Score 261.4; DB 20; Length 2376;
Best Local Similarity 77.6%; Pred. No. 2e-45;
Matches 316; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
OY 65 CCGGCGGCGCGCGCGCACTCCACCGCGCGCGCGCGCGCAACCAAGAAACAGCC 124
D 518 CCGCGCGGCGCGCGCGGTGGGGCGCGCGCGCGCGCGCGCGCGCGCAACCAAGCC 577
OY 125 CGAGCGCGCTCAAGCGCGCGCATGAATGCTTCATGATGTGCTCCCGGGGCGGCGCA 184
D 578 AGGATGCGGTCAAGCGCGCGCATGAATGCTTCATGATGTGCTCCCGGGGCGGCGCA 637
OY 185 AGATGCGCCAGAGAAACCCCAAGATGCACAATCGAGATCAGCAACGCTGGCGCGC 244
D 638 AGATGCGCCAGAGAAACCCCAAGATGCACAATCGAGATCAGCAACGCTGGCGCGC 697
OY 245 AGTGAACCTTTTGTGCGAGACGAGAGCGCGCGCTTCATCAGCAGGCTTAAGCGGCTGC 304

```


XX WPI: 2001-457597/49.
DR p-PSDB: AAE05814.
XX
PT Isolated polypeptide, used to treat or prognose a disorder
PT characterized by expression of a hCAP e.g. cancer, is encoded by an
PT isolated nucleic acid comprising an NA Group 3 or 4 molecule -
XX
PS Claim 57, Page 100-105, 152pp; English.

XX The invention relates to nucleic acids and encoded polypeptides which
CC are cancer associated antigens expressed in patients afflicted with
CC small cell lung cancer. The molecules provided by the invention can be
CC used in the diagnosis, monitoring, research or treatment of conditions
CC characterised by the expression of one or more cancer associated
CC antigens. The polypeptide is used to treat a disorder characterised by
CC expression of a hCAP, and determine regression, progression or onset
CC of a condition characterised by expression of an abnormal amount of a
CC protein encoded by a nucleic acid (NA) Group 1 molecule. The disorders
CC are small and non-small cell lung cancer, melanoma, colon, breast, head
CC and neck, or transitional cancer, leiomyosarcoma or synovial sarcoma.
CC The present sequence is a small cell lung cancer associated gene
CC designated as NY-SCLC-10, encoding human SOX21 protein.
XX

XX Sequence 8372 BP; 1996 A; 2143 C; 2104 G; 2124 T; 5 other;

XX Query Match 45.0%; Score 225; DB 22; Length 8372;
XX Best Local Similarity 71.6%; Pred. No. 7.5e-38;

XX Matches 310; Conservative 0; Mismatches 120; Indels 3; Gaps 1;

QY 67 GGGGGGGGGGGGGGCAACTCACCAGGGGGGGGGGGGGGCAACCAAGAAACAGCCCG 126
DB 1122 GAGGG 1181
QY 127 GACCGGCTCAAGGG 186
DB 1182 GACGACAGTCAAGGG 1241
QY 187 ATGGCCCAAGAGAACCCCAAGATGACACACTCGAGATCAAGCAAGCGCTTGGGGCGAG 246
DB 1242 ATGGCCCAAGAGAACCCCAAGATGACACACTCGAGATCAAGCAAGCGCTTGGGGCGAG 1301
QY 247 TGGAAACTTTTGGGAGAGCGGAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 306
DB 1302 TGGAAACTTCTCAAGAGTGGGAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1361
QY 307 GCGCTGCACATGAAGAGCAACCGGATTTATTAATATCCGGGGGGGGGGGGGGGGGGGG 366
DB 1362 GCCATGCACATGAAGAGCAACCGGATTTATTAATATCCGGGGGGGGGGGGGGGGGGGG 1421
QY 367 CTCATGAAGAAGATTAAGTACAGCTGCGCG---GCGGGCTGCTGGCCCGCGCGCAAT 423
DB 1422 CTCCTCAAGAAAGAGCAAGTTCGCTTCCTCCGTCCTTACGCGGGGGGGGGGGGGGGGG 1481
QY 424 ACCATGGCAGAGCGGGGTGCGGGGTGGGGCGGCTGGGGCGGGCGGAAACCAAGCATG 483
DB 1482 GCCGAGCAACCTTGCGGTCAAGGGGGGGGGCGCGGGCTGACCGGGGGGGGGGGGGGGCT 1541
QY 484 GACAGTTACGGCG 496
DB 1542 GTGCTGAGTGC 1554

RESULT 14

XX AAS87738 standard; cDNA; 2378 BP.

XX AAS87738;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #23542.

KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YF;

XX WPI: 2001-639362/73.

XX p-PSDB: ABG23551.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX Claim 1; SEQ ID No 23542; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

XX Sequence 2378 BP; 590 A; 708 C; 664 G; 416 T; 0 other;

XX Query Match 44.0%; Score 220; DB 23; Length 2378;
XX Best Local Similarity 69.0%; Pred. No. 7.5e-37;

XX Matches 334; Conservative 0; Mismatches 140; Indels 10; Gaps 2;

QY 27 GGAGCGAGCTGAAGCCGCGGGCCGCAACAACTTGGGGGGGGGGGGGGGGGGGGCAACTC 86
DB 733 GGAGCGCGAGCGCAAGATGATGCGAAGCGAGCGGGGGGGGGGCAACTTGGGGGGGGGG 792

QY 87 CACCGGG 145
DB 793 AGCGGTGTGGAGCGAGGTGGCGGGGTACAGACCAAGACCGGTGTAACCGGCCCAA 852

QY 146 TGAATGCTTCATGATGTGTGTCGCCGCGGACAGCGCGCAAGATGGCCCGAGGAACCCCA 205
DB 853 TGAAGCCTTCATGATGTGTGTCGCCGCGGACAGCGCGCAAAATGGCCCGAGGAACCCCA 912

QY 206 AGATGCACACTTCGAGATTCAGCAAGCGCTGGGGCGCGAGTGAACCTTTTGTGGAGA 265
DB 913 AGATGCACAAATTCGAGATTCAGCAAGCGCTGGGGCGCGACTGGAACACTGTGACGACAG 972

QY 266 CGGAGAGCGGGCGTTTCATTCAGACGAGGCTAAGCGGCTGGAGCGCTGCACATGAAGAGC 325

```

Db 973 CCGAAGAGCGACCATTCATCGACGAGCCAGACGACTTGCCGCCCTGCACATGAAGAGT 1032
QY 326 ACCGGGATTATTAATATACGGGCCCCCGGGAACCAAGACGCTCATGAAGAAGATTAAGT 385
Db 1033 ATCCGGACTTACAAAGTACCGACCGCCGCAAGACCAAGACGCTGCTCAAGAAAGATTAAGT 1092
QY 386 ACACGCTGCCGCGGCGGCTGCTGCGCCCGCGGCAATGACATGAGCGAGCGGGGTGGG- 444
Db 1093 ACTGCTGCTCCGACGGGCTCTCTGCTCCCGGCTGCGCGGCGCCGCCGCTGCGCGG 1152
QY 445 -----GTGGGCGCGCGGCTGCGGCGCGGCTGAACCGCATGACATTAACGCGC 496
Db 1153 CCGGAGCGCGTGCAGCGCAGAGTCCGCTGGCGGCGCGCTGAGACAGTACACGC 1212
QY 497 ACAT 500
Db 1213 AGCT 1216

```

RESULT 15

AAS61625
ID AAS61625 standard; cDNA: 412 BP.

XX AC AAS61625;

XX DT 29-JAN-2002 (first entry)

XX DE Lung small cell carcinoma antigen, cDNA #166.

XX KW Human; cytosolic; antitumour; lung small cell cancer antigen;

XX KM tumour; lung cancer; ss.

XX OS Homo sapiens.

XX PN W0200177168-A2.

XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001W0-US11859.

XX PR 11-APR-2000; 2000US-196780P.

XX PR 21-JUN-2000; 2000US-213361P.

XX PR 01-SEP-2000; 2000US-229763P.

XX PR 05-SEP-2000; 2000US-230629P.

XX PR 14-SEP-2000; 2000US-232565P.

XX PR 19-DEC-2000; 2000US-257037P.

XX PR 08-JAN-2001; 2001US-260796P.

XX PA (CORI-) CORIXA CORP.

XX PI Lodes MJ, Wang T, Mohamath R, Indirias CY;

XX DR WPI; 2002-010896/01.

XX PT Lung tumour polynucleotide and polypeptides useful in therapy and

XX PT diagnosis of cancer especially lung cancer

XX PS Claim 1; Page 180; 295pp; English.

XX CC The invention relates to novel isolated lung small cell cancer antigen
 CC polynucleotides (I) and polypeptides (II) used in a method of detecting
 CC cancer in a patient. The method is optionally performed by
 CC utilising oligonucleotides (III), where the biological sample
 CC from the patient is contacted with (III), detecting the amount of
 CC polynucleotide hybridised to (III) in the sample and comparing the
 CC amount of polynucleotide to a predetermined cut-off value and thereby
 CC determining cancer in a patient. (I), (II) or antigen-presenting cells
 CC expressing (II) is useful for stimulating and/or expanding T cells
 CC specific for a tumour protein. The method comprises contacting T cells
 CC with one of the components under conditions to permit the stimulation
 CC and/or expansion of the cells. A composition comprising (I) is useful for
 CC stimulating an immune response in a patient and for inhibiting the

CC development of a cancer especially lung cancer in a patient. An
 CC isolated T cell population is useful for removing tumour cells from the
 CC biological sample and for inhibiting the development of cancer in a
 CC patient. AAS61460-AAS61874 represent novel human lung small cell
 CC cancer antigen coding sequences of the invention.

SQ Sequence 412 BP; 83 A; 133 C; 144 G; 51 T; 1 other;

Query Match 43.6%; Score 217.8; DB 24; Length 412;

Best Local Similarity 75.7%; Pred. No. 1.9e-36;

Matches 283; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

```

QY 126 GGACGGGCTCAAGCGGCGCCCATGAATGCTTCATGCTGTGTGTCGCGGCGGCAAGCA 185
Db 15 GGACGAGTCAAGCGGCGCCCATGAATGCTTCATGCTGTGTGTCGCGGCGGCAAGCA 74
QY 186 GATGGCCGAGGAACCCCAAGATGACACACTGCGAGATCAGCAAGCGCTTGGGCGCA 245
Db 75 GATGGCCGAGGAACCCCAAGATGACACACTGCGAGATCAGCAAGCGCTTGGGCGCA 134
QY 246 GTGGAACCTTTGTCGAGAGCGAGAACGCGCCCTTCATCCAGCGCTAACGCGCTGCG 305
Db 135 GTGGAACCTGCTACAGAGTGGAGAACGCGCCCTTCATCCAGCGCGCCACGCTTACG 194
QY 306 AGCGCTGCACATGAAGGAGCACCGGATTTAATACCGGCGCGGGAACCAAGAC 365
Db 195 CGCCATGCACATGAAGGAGCACCGGATTTAATACCGGCGCGGGAACCAAGAC 254
QY 366 GCTCATGAAGAGATTAATGACAGCTGCCG--GCGGGCTGCTGCGCCCGCGGCGCA 422
Db 255 GCTCATGAAGAGATTAATGACAGCTGCCGCTTCAGGCGCTGAGGCGGCGGCGCA 314
QY 423 TAGCATGGCGAGCGGGGCTGGGCGCGGCGCTGGGCGGCGGCGGCGGCGGCGCAT 482
Db 315 CGCGGAGCACCTGCGCTCAAGGCGGCGGCGGCTGACCGGCGGCGGCGGCGGCGCT 374
QY 483 GACAGTTACGCGC 496
Db 375 GGTGCTGAGTCCG 388

```

Search completed: October 10, 2002, 14:28:33
 Job time : 92 secs

